TOWARDS A DEFINITION OF SCHWA: AN ACOUSTIC INVESTIGATION OF VOWEL REDUCTION IN ENGLISH

Sally Alexandra Rosemary Bates

PhD
University of Edinburgh
1995

I declare that the production of this thesis
and the work reported herein were carried out by me
except where otherwise stated

Acknowledgment

There are many people who have contributed to the preparation of this thesis and who I would like to thank. My first debt is to my mother. Thank you for all the support, understanding and encouragement you have shown me over the years and for enabling me to pursue my chosen interests. I also acknowledge my father and his support and encouragement when I first went to university as an undergraduate. He would have been so pleased and so proud to see me complete this work. Also, deserving mention for being important people are Duncan, Ginny and Pamsie.

I would also like to express my gratitude to those people who have given me invaluable guidance during the course of this research. I am especially grateful to my supervisor Stephen Isard for his wise and insightful advice and for his patience and encouragement. Thanks are also due to Ellen Bard for helping me not only to cope with, but also to enjoy, the statistical analysis of my data. I would also like to thank Jonathan Harrington, Gerry Docherty and Jim Hieronymous for their advice and guidance in the early stages of this research. Thank you also to Alan Kemp and Jim Scobbie for reading parts of this thesis and for their valuable comments.

I am also grateful to all those who have given their technical assistance. Thank you to the technical and computing staff at the Department of Linguistics and CSTR. Special thanks are due to Paul Bagshaw for the time and application he has given to the various projects I have sought assistance with. Thank you also to Gordon Watson and Art Blokland for their help with S and APS and to Fergus MacInnes for sharing his mathematical expertise with me. I am also indebted to the School of Computer Science at Plymouth University for allowing me to use their facilities during the final stages of this work.

Other people in the department who deserve special mention include Ethel Jack and Irene MacLeod. Their friendship and support have been invaluable. Thank you also to all those in my peer group, in particular Sandra Vella, Louise Kelly, Yuko Kondo, Robin Lickley, Mark Schmidt and Mike Moroney. They have provided interesting discussion and have also helped make this a very enjoyable time. An extra special thank you is due to my friend and colleague Jocelynne Watson. Thank you Jocelynne for your critical eye and for your humour. I couldn't have managed without you.

The last word of acknowledgement and dedication of this thesis goes to my husband, Phil Shingler. I cannot begin to do justice to the gratitude I feel for the support he has given me. Our first few months of married life has coincided with the writing up of this research. Thank you Phil for your tolerance and understanding during this time, you have been wonderful.

Abstract

This thesis reports a single speaker acoustic study of vowel variability in connected speech. Over eight thousand vowel tokens taken from a corpus of read sentences are examined. The aim of the thesis is to achieve a better understanding of the nature of vowel reduction in English. Three questions are addressed. The first of these concerns the phonetic characterisation of schwa, the central or 'reduced' vowel. Schwa's contextual variability is assessed with reference to the question of whether or not it has an independent phonetic target. The second question concerns the role of stress in conditioning vowel reduction. Patterns of variability for sententially stressed and unstressed vowel tokens are examined in order to determine how far stress and context effects interact to influence vowel quality. The final question concerns the potential differences between vowels with respect to inherent variability, that is, whether some vowels are inherently more susceptible to coarticulatory effects than other vowels.

Maximal context-dependency for schwa strongly supports the hypothesis that it is completely unspecified for tongue position. The data indicate that it is also highly unspecified for jaw position. Evidence that schwa is targetless and can occupy almost any position in the vowel space depending on context, argues against the traditional concept of vowel reduction as an independent process of articulatory and/or acoustic centralisation. Greater context sensitivity for sententially unstressed vowels compared with their sententially stressed counterparts also supports an account of vowel reduction in terms of contextual assimilation.

The results also indicate a continuum of underspecification. This ranges from the more peripheral vowels /i, a, ɑ, ɒ, ɔ/ which show the least contextual variability and which may be thought of as the most narrowly specified vowels to the more central vowels /ɪ, ɛ, ɜ, ʌ, ʊ/ and, in the present data, /u/, which show greater overall context-dependency. It is proposed that greater acoustic stability for the more peripheral vowels reflects quantal acoustic properties and tighter articulatory and/or perceptual constraints on variability.

Overall, the results support the view that vowel reduction represents a means of economising on articulatory effort. Schwa, the endpoint of the reduction process represents minimal articulatory effort insofar as it represents the straight-line interpolation between consonants and hence minimal resistance to coarticulatory effects. Shorter durations, greater context dependency and, in the case of the peripheral vowels, less extreme formant values for sententially unstressed compared with sententially stressed vowels reflects a reduction in articulatory effort and consequently less displacement from neutral. In view of the greater context-dependency observed for the more central vowels generally compared with the more peripheral vowels, the tense/lax alternation in phonological vowel reduction can also be interpreted as a saving on articulatory effort. A principal advantage of an account of English vowel reduction in terms of phonetic underspecification is that phonetic and phonological vowel reduction may be accounted for by the same mechanism.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

The aim of this thesis is to achieve a better understanding of the nature of vowel reduction in English. Three related questions which have a direct bearing on this issue are addressed. The first question concerns the phonetic characterisation of the central or 'reduced' vowel /ə/, also referred to as 'schwa'. Schwa's contextual variability is assessed in the light of the proposal that it has no independent phonetic target but is completely unspecified for tongue and jaw position. The second question concerns the role of context and stress in conditioning vowel reduction. Patterns of variability displayed by sententially stressed and unstressed full vowel tokens are examined in order to determine how far stress and context effects interact to influence vowel quality. The third question concerns the potential differences between vowels with respect to inherent variability, that is, whether some vowels are inherently more susceptible to coarticulatory effects than other vowels.

1.1 Vowel reduction

Vowel reduction refers to the alternation of full vowels with schwa in unstressed syllables and is a characteristic feature of stress-timed languages such as English. A more detailed description is given in Chapter 2. The precise nature of vowel reduction, however, remains to be established. Traditionally, it is conceived of as an independent process of centralisation towards the neutral vowel /ə/ (Tiffany, 1959; Shearme & Holmes, 1962; Delattre, 1969; Stalhammer et al., 1973; Koopmans van Beinum, 1980; Fourakis, 1991 inter alia.). According to this view, unstressed vowels are generally more central in quality than stressed vowels, that is, they are produced with less vocal effort resulting in less extreme articulatory configurations and less distinctive acoustic patterns.

An alternative account proposed by Lindblom (1963) is that reduction in vowel quality reflects target undershoot due to coarticulation. According to Lindblom, unstressed vowels are more susceptible to coarticulatory effects than stressed vowels because they are shorter in duration. Stressed and unstressed vowels are characterised by the same invariant target configuration. However, in producing an unstressed vowel token, the speaker undershoots the target due to the temporal overlap of motor commands and articulatory "sluggishness". A more detailed discussion of Lindblom's undershoot hypothesis and its theoretical implications is given in section 2.3.1 and in the introduction to Chapter 6.

In subsequent research following Lindblom (1963), it has been shown that vowel quality may vary independently of duration (Gay, 1978; Tuller, Harris & Kelso, 1982; Nord, 1986; Summers, 1987; Den Os, 1988; Engstrand, 1988; Van Son, 1993 inter alia.). The results from these studies indicate that the speaker is able to manipulate coarticulatory parameters which, in turn, suggests that reduction forms part of the speaker's stylistic intent. In view of this, Lindblom (1983, 1990) adapts the original undershoot hypothesis to include a mechanism whereby speakers are able to override or avoid coarticulatory effects through increased vocal effort and/or a reorganisation of phonetic gestures. However, reduction in vowel quality is still attributed to durationally-induced target undershoot.

In this thesis, evidence is sought in support of an account of vowel reduction which reconciles elements of both the traditional approach and Lindblom's model. In line with Lindblom's model, it is proposed here that vowel reduction may be accounted for in terms of contextual assimilation[1]. However, in contrast to the undershoot hypothesis, it is suggested that unstressed vowels are more susceptible to coarticulatory effects than stressed vowels because they are inherently less narrowly specified. That it, they are produced with less articulatory effort resulting in less

[1] It is noted here that 'contextual assimilation' is used to refer to coarticulatory adjustment rather than phonological feature-spreading.

displacement from neutral where 'neutral' is defined either in terms of the overall rest position of the vocal tract or the straight-line interpolation between adjacent context segments. This account thus accords with the traditional approach insofar as it posits a separate specification for stressed and unstressed vowels.

The suggestion that unstressed vowels are more susceptible to coarticulation because they are inherently less narrowly specified than stressed vowels is made in accordance with Keating's (1988, 1990) theory of phonetic underspecification. Keating (1988) proposes that segments which are inherently unspecified for a given feature remain unspecified at the phonetic level and are simply "interpolated through" by the trajectory between adjacent specified segments. Keating (1990) further proposes that segments may show varying degrees of underspecification for a given feature. At the phonetic level, this manifests itself in varying degrees of contextual variability along the relevant dimension(s). Segments are thus characterised by the full range of contextual variability they exhibit rather than by a single set of target spatial coordinates. A more detailed description of the underspecification hypothesis is given in sections 2.3.2.2 and 2.5.

In the current investigation, sententially stressed and unstressed vowel tokens are compared in terms of the magnitude and patterns of contextual variability they exhibit. Assuming that unstressed vowels are less narrowly specified for tongue-body and jaw position than stressed vowels, they are expected to show greater overall variability and context-dependency along the corresponding acoustic dimensions F2 and F1.

In practise, it may be difficult to differentiate between Lindblom's undershoot hypothesis and the proposed phonetic underspecification account of vowel reduction since both predict increased contextual assimilation for unstressed vowels relative to stressed vowels. However, additional support for an underspecification account may be provided by observing patterns of variability across vowels as well as for stressed and unstressed tokens within a given vowel category. Evidence of inherent differences between vowels with respect to the contextual variability they exhibit would be consistent with Keating's proposal that segments may be more or less fully specified along a given dimension(s), phonologically and phonetically.

1.2 Inherent vowel variability

There is an increasing body of literature which reports greater context sensitivity for unstressed vowels relative to stressed vowels (Lindblom, 1963; Nord, 1974, 1986; Fowler, 1981; Magen, 1984, 1989; Krull, 1989; de Jong et al., 1994). However, comparatively little attention has been given to the question of whether context and/or stress affects all vowels to the same degree.

Assuming that unstressed vowels are less resistant to coarticulation than stressed vowels because they are articulated with less vocal effort and less divergence from neutral, similar differences in degree of contextual variability might be expected for the more central vowels generally (i.e. /ɪ, ɛ, ɜ, ə, ʌ, ʊ/), compared with the more peripheral vowels (i.e. /i, a, ɑ, ɒ, ɔ, u/). Stevens (1972, 1989) quantal theory of speech production predicts greater acoustic stability for the point vowels /i, ɑ, u/ compared with the non-point vowels (see sections 2.6 and 7.2.3). However, to date, there has been relatively little research directed towards establishing empirical support for quantal theory predictions.

In an electropalatographic and acoustic study of consonant-to-vowel coarticulation, Recasens (1991) reports greater variability for back vowels than for front vowels and greater variability for schwa than for the full vowels. Recasens proposes that front vowels are inherently more stable than back vowels because they involve a greater degree of mechanical constraint on the tongue body during their production. Conversely, he attributes the high levels of variability observed for schwa to low constraints on its tongue-body specification. Recasens' articulatory constraint based model of coarticulation is discussed in more detail in section 2.3.2.1.

Recasens' observations are based on nonsense speech data for vowels in Catalan. The main focus of his investigation is electropalatographic data although he also reports acoustic data. A principle aim of this thesis is to further explore Recasens' and Stevens' (1989) predictions regarding the relative variability of point/non-point and front/back vowels in an acoustic investigation of connected speech data in English.

1.3 Variability of schwa

In line with the phonetic underspecification hypothesis (Keating, 1988, 1990), Browman & Goldstein (1992) suggest that schwa may be completely unspecified for tongue position. However, following their articulatory study of schwa tokens produced in /pVpə'pVp/ sequences, they reject this idea and argue instead in favour of a co-production account of schwa's variability. A full description of their study is given in section 2.3.2.2. Evidence in support of the targetless analysis, however, has been presented in a recent paper which reports data for Dutch schwa in a wider range of contexts (Van Bergem, 1994). Van Bergem reports highly systematic coarticulatory effects on schwa such that the schwa second formant values can be predicted using a simple linear model. Van Bergem's study is decribed in more detail in section 2.4.

The present study represents an extension of Browman & Goldstein's (1992) and Van Bergem's (1994) research insofar as it examines coarticulatory effects for schwa in connected speech data and in a more comprehensive range of contexts. As in the case of Browman & Goldstein, Van Bergem's observations are largely based on nonsense data although he found that his model worked well on a set of a hundred meaningful words.

The present investigation also provides a comparative framework in which to assess schwa's variability. Without reference to comparative data for the full vowels, it is difficult to assess how much context-dependency may be considered to reflect targetlessness. Van Bergem (1994) reports 72% and 79% explained variance in schwa F2 midpoint value for schwa tokens in nonsense syllables. Given that laboratory speech may undersample the phonetic variability found in more natural speech data, it is possible that the extent of the contextual variation shown by the full vowels in real speech may approach or even exceed the levels of variability reported for schwa in laboratory speech. In absolute terms, the proportion of explained variance obtained in regression analyses/analyses of variance, is also a fairly arbitrary measure due to the fact that it is subject to variation across different samples of the same dependent variable (see Cohen & Cohen, 1983, p.5). In view of both these factors, a comparative study using the proportion of explained variance as a relational measure permits a more reliable evaluation of the data for schwa and hence a more conclusive assessment of schwa's targeted/targetless status.

1.4 Connected speech data

To date there have been no studies of schwa's variability in which schwa is compared with the other monophthongs in a British English accent system. Furthermore, there is relatively little quantitative data available for either schwa or the full vowels for connected speech data. Most of the research on vowel-to-vowel or consonant-to-vowel coarticulatory effects examines nonsense syllables or isolated words typically produced in a semantically empty carrier phrase. These are also carefully controlled with respect to context and stress and tend to be limited to a restricted set of vowels and consonants. While the controlled, experimental paradigm is without doubt fundamental to speech production research, there is also a recognised need for research based on more natural speech data (Rischel, 1990).

A number of studies indicate that the magnitude and patterns of coarticulation vary considerably across speech styles (Koopmans van Beinum, 1980; Krull, 1989; Farnetani et al., 1993; Harmegenies & Poch-Olive, 1992). Therefore, the extent to which observations based on laboratory speech extend to connected speech data clearly has important implications for the generalisability of proposed models of coarticulation. Observational studies of connected speech data also have an important role to play in establishing the limits of phonetic variability. In addition, given that reduction processes are conditioned by semantic and pragmatic as well as phonetic variables, it is also important that these should be examined in more natural speech data which reflects the interaction of these variables (Lindblom, 1990).

The speech material used in the present study comprises six hundred and sixty read sentences. While sentences read under experimental conditions do not constitute spontaneous, conversational data, they nevertheless offer a means of examining fluent, meaningful speech data. In this respect, they may be considered a useful transitional step between the study of highly controlled laboratory speech and the study of spontaneous speech.

In more general terms, the nature and extent of phonetic variability also has important implications for the invariance issue in speech production. For example, evidence of systematic acoustic variability would support the argument in favour of signal-based invariance (Stevens & Blumstein, 1979, 1981; Blumstein & Stevens, 1979; Sussman, 1990b, 1994; Sussman et al., 1991) over gestural-based theories (Liberman & Mattingly, 1985; Fowler, 1986).

1.5 The structure of the thesis

Chapter 2 gives the theoretical background to the thesis. It provides a more detailed definition of vowel reduction and explores alternative predictions with reference to the question of inherent vowel variability. It also describes the underspecification hypothesis in detail and reviews the two studies to date which specifically address the question of whether or not schwa can be associated with an independent phonetic target.

Chapter 3 describes the speech material and data-analytic procedures that are used in this study. It also provides some preliminary acoustic data to orientate the reader with respect to the current subject's vowel system.

Chapter 4 addresses the methodological problem of how to represent the spectral vowel target. In the first part of the chapter, the classificatory performance of five of the most commonly employed sampling points and single number representations of vowel formant frequencies are compared and assessed in terms of the statistical separability of the derived vowel distributions. In the second part, Syrdal & Gopal's

(1986) bark-difference model of vowel recognition is evaluated with reference to the question of whether bark-transformed spectral distance measures offer any advantage over linear spectral distance measures or absolute frequency values in normalising within-speaker contextual variability.

Chapters 5 and 6 examine the data with respect to the question of whether some vowels are inherently more stable than others. Chapter 5 investigates the role of context in conditioning vowel quality. The relative magnitude and pattern of coarticulatory effects for different vowels are assessed statistically in order to establish a hierarchy of vowel robustness. The data is also examined with respect to the question of schwa's phonetic status. Chapter 6 considers the influence of stress and the interaction of stress and context effects. It provides a review of the relevant literature and an analysis of the comparative variability displayed by sententially stressed and unstressed vowel tokens in the present data. It also considers the question of whether sentence stress influences vowels to the same extent or whether some vowels show relatively greater stress effects than others.

The final chapter (Chapter 7) considers the results reported in the analysis chapters with reference to a phonetic underspecification account of vowel reduction. It also offers an interpretation of the results in the light of predictions made in the literature with respect to production and perceptual constraints on variability. The theoretical implications of the extent and nature of the observed variability for theories which advocate signal-based invariance are also discussed. The chapter concludes with a summary of the main points arising from the thesis.

Chapter 2
Inherent Vowel Variability

This chapter establishes the theoretical framework for the thesis. It introduces the concepts of coarticulation and phonetic vowel reduction and explores alternative accounts for inherent differences between vowels with respect to the phonetic variability they exhibit. Phonological vowel reduction is also described and the relationship between phonological and phonetic vowel reduction discussed.

2.1 Coarticulation

Variability in the articulatory and acoustic realisation of segments is a characteristic feature of natural, continuous speech. Random variability occurs because the talker is unable to exactly reproduce a given utterance on any two occasions. Systematic variability occurs as a function of phonetic context. Traditionally, contextual variability is viewed as the deviation or displacement from idealised, target values (Shearme & Holmes, 1962; Stevens & House, 1963; Lindblom, 1963; Stevens, House & Paul, 1966; Henke, 1966; MacNeilage & DeClerk, 1969; Daniloff & Hammarberg, 1973; Ohde & Sharf, 1975 inter alia.). Segmental targets are equated with the articulatory configuration adopted, and the corresponding acoustic pattern obtained, when the segment is produced in isolation. Because the vocal tract cannot change instantaneously from one configuration to another, interactive influences in the articulatory and acoustic transitions between segments are inevitable when they are produced in sequence. The process posited to account for the interaction between segments in connected speech is coarticulation.

Coarticulation is bi-directional in that it may operate from left-to-right and/or from right-to-left. Left-to-right or "carry-over" coarticulation refers to the influence of a preceding segment on a segment currently in production. Right-to-left or "anticipatory" coarticulation obtains when a segment currently in production is influenced by the production requirements for a following segment. Carry-over effects are thought to be largely mechano-inertial in nature while anticipatory effects are believed to be timing effects, reflective of articulatory pre-programming (Gay, 1977).

Coarticulatory effects extend from consonants to vowels (Lindblom, 1963; Stevens & House, 1963; Stevens, House & Paul, 1966; Ohde & Sharf, 1975; Recasens, 1991), from vowels to consonants (Kozhevnikov & Chistovich, 1965; Moll & Daniloff, 1971; Benguerel & Cowan, 1974; Kent, Carney & Severeid, 1974; Schouton & Pols, 1979; Recasens, 1984, 1986), from vowels to vowels (Ohman, 1966, 1967; Kent & Moll, 1972; Bell-Berti & Harris, 1976; Gay, 1977; Fowler, 1981; Manuel & Krakow, 1984; Recasens, 1986; Magen, 1984, 1989; Manuel, 1990) and from consonants to consonants (Byrd, 1994). The range and relative strength of these different effects have important implications for the proposed nature of the underlying units of speech production.

2.2 Phonetic vowel reduction

In the case of vowels, coarticulation is generally viewed as a process of "feature-reduction" (Sharf & Ohde, 1981). Vowels in context are typically more central or reduced in quality than equivalent vowels produced in isolation. The influence of adjacent consonants tends to lower F2 in the case of front vowels and raise F2 in the case of back vowels (Stevens & House, 1963). However, as Lindblom (1963) points out, failure to attain target formant frequency values as a function of context does not necessarily imply articulatory and acoustic centralisation. Although a global reduction in the degree of acoustic contrast between vowels may be observed for vowels in context compared with vowels in isolation (Koopmans van-Beinum, 1980), at the level of the individual vowel token, the shift in formant frequency is not necessarily towards the center of the vowel space. Rather, the direction and the extent of the shift will depend on the vowel in question and the context in which it occurs. For example, in the context of bilabial consonants which are associated with low frequency loci (Delattre et al., 1952), undershoot for the low, front vowel /a/ would be expected to result in an F2 frequency value closer to the putative neutral value of 1500 Hz than to its target value in the region of 1750 Hz (Wells, 1962). However, in the context of consonants characterised by high F2 locus frequencies such as palatals and velars, failure to attain the vowel target would result in a value that is further from, rather than nearer, the neutral value.

It is therefore important, when describing coarticulatory effects for vowels, to clarify the definition of terms such as "vowel reduction", "neutralisation", "target undershoot" and "centralisation", all of which tend to be used interchangeably in the literature (see Tokuma, 1993 for a review). In this thesis, “phonetic vowel reduction” is used to refer to the contextual variation of vowels. It is assumed that, while phonetic vowel reduction often results in more central values for vowels compared with ideal, target values, the displacement in vowel formant frequency is not necessarily towards the neutral value traditionally associated with schwa. Phonetic reduction is preferred over “target undershoot” because, unlike the latter term, it is not allied to a particular model of coarticulation. A distinction is also assumed between phonetic and phonological vowel reduction which refers to the lexical alternation of full vowels with schwa (see section 2.7).

2.3 Inherent differences between vowels with respect to degree of context-sensitivity

There is evidence in the literature to suggest that vowels vary inherently with regard to the extent to which they are susceptible to coarticulatory effects. For example, greater contextual variability has been observed for lax vowels compared with tense vowels (Stevens & House, 1963; Stevens, House & Paul, 1966; Strange, 1989a), for back vowels compared with front vowels (Syrdal & Gopal, 1986; Sussman, 1990a; Recasens, 1991) and for schwa compared with the full vowels (Flege, 1988;

Recasens, 1991; Koopmans van Beinum, 1994). Unstressed vowels have also been shown to be more sensitive to context than stressed vowels (Lindblom, 1963; Nord, 1974, 1986; Fowler, 1981; Magen, 1984, 1989; Krull, 1989; de Jong et al., 1994).

Hypotheses regarding the inherent variability of vowels have been formulated on the basis of articulatory, acoustic and perceptual considerations. In terms of articulatory constraints on variability, there are essentially two accounts for the observed differences between vowels with respect to the amount of contextual variation they exhibit. The first of these follows the traditional target undershoot approach to coarticulation where degree of coarticulation is dependent upon the distance between, and the amount of time available in which to reach, segmental targets (Stevens & House, 1963; Lindblom, 1963). The second follows the more recent gestural approach to coarticulation, where coarticulation is regulated by the degree of articulatory compatibility between segments (Recasens, 1984, 1986, 1991; Browman & Goldstein, 1984, 1986, 1989; Boyce et al., 1990).

2.3.1 Target undershoot approach to coarticulation

According to the target undershoot approach to coarticulation (Stevens & House, 1963; Lindblom, 1963), phonemes are represented in the speaker's phonetic plan as a successive sequence of discrete, invariant target configurations. The articulators move toward these targets in response to neural commands:

> During the production of the initial consonant of a syllable, the structures are first caused to assume a configuration that is characterised by a vocal-tract constriction at a point that depends on the consonant's place of production. The structures are then instructed to maneuver toward a configuration that is appropriate to the vowel. During the course of this maneuver, a new instruction indicates what the articulatory configuration for the final consonant should be, and the structures move without discontinuity toward this configuration. (Stevens & House, 1963, p.122)

Because the vocal apparatus as a physical system is characterised by the properties of mass and inertia, the success with which a given target is attained depends to a large extent upon "the effective 'distance' that the articulators must traverse during the various phases of the syllable between the initial consonant, the central syllabic nucleus and the terminal consonant" (p.120). It also depends upon the amount of time there is available for the articulators to reach the configuration specified by one set of motor commands before they have to start moving in response to the next set of commands.

According to Lindblom (1963), unstressed vowels and vowels produced at a fast rate of speech are more susceptible to coarticulatory effects than their stressed or more slowly produced counterparts because they are shorter in duration. Shorter durations result in greater overlap in the timing of commands to the articulators which, in turn, result in a greater degree of target undershoot at both the articulatory and acoustic level (see Chapter 6, Introduction for a full discussion).

Assuming that reduction is duration dependent, differences in inherent variability between tense and lax vowels and between schwa and the full vowels might also be attributed to durational differences. With the exception of /a/, lax vowels are inherently shorter in duration than tense vowels (Peterson & Lehiste, 1960). Of all the vowels, schwa has the shortest intrinsic duration.

The majority of studies following Lindblom (1963) which investigate the relationship between undershoot and duration, examine changes in vowel quality that occur as a function of variations in duration due to stress and/or speech rate (Tuller et al., 1982; Harris, 1968; 1971, 1973; Gay et al., 1974; Gay, 1974, 1977, 1978; Verbrugge & Shankweiler, 1977; Sussman, 1978; Fourakis, 1991) or phonetic context (Summers, 1987). Krull (1989) includes phonological length as a factor in her analysis of consonant-vowel coarticulation in Swedish. However, she reports little difference in amount of coarticulation for long compared with short vowels. Greater differences are observed for stressed/unstressed pairs of vowels. This suggests that it is stress and not duration which determines reduction. (The relationship between duration and reduction is discussed more fully in Chapter 6.) However, it is also possible that the phonological length distinction in Swedish does not have the same implications at the phonetic level as the tense/lax distinction in English. That is, that properties other than intrinsic duration are responsible for differences in degree of context-sensitivity between tense and lax vowels.

There is evidence that tense and lax vowels in English are also distinguished by differences in the temporal structure of the formant trajectories. Lax vowels are characterised by relatively shorter onglide durations, shorter steady-state portions and longer offglide durations proportional to the total vowel duration than tense vowels (Stevens & House, 1963; Stevens, House & Paul, 1966; Lehiste & Peterson, 1961; Strange, 1989a). They also differ with respect to the direction of formant movement away from the target. Stevens & House (1963) report that for lax vowels, "during the remainder of the syllabic nucleus the configuration moves toward a neutral position" (p.130). Nearey (1989) observes a similar shift in formant frequency value for the lax vowels /ɪ/, /ɛ/ and /a/ even when produced in isolation. In contrast, the formant frequencies for tense vowels move towards the extremes of the vowel space. This distinction is also noted by Strange (1989a). The greater variability of lax vowels compared with tense vowels might therefore be attributable to the fact that they are articulated with less vocal effort and less divergence from neutral.

2.3.2 Gestural approach to coarticulation

The gestural approach to coarticulation provides a framework in which the greater variability of lax compared with tense vowels, schwa compared with the full vowels and unstressed compared with stressed vowels may be accounted for in terms of reduced vocal effort rather than a reduction in duration per se.

In gestural models, segments are described indirectly in terms of synchronised gestures or functional groupings of gestures rather than as spatial targets. Coarticulation is attributed to the overlapping production of these gestures. Individual gestural models vary with respect to the question of how the production units of speech are defined and how precisely phonetic structure is implemented. However, in general terms, degree of coarticulation is considered to be dependent upon the extent to which the gestures required for a given sequence of segments are complementary or antagonistic to each other. In line with this approach, more central vowels arguably show greater susceptibility to coarticulation than more peripheral vowels because their articulatory requirements are less likely to conflict with those of other segments.

2.3.2.1 Recasens (1986, 1991) articulatory constraint-based model

Recasens (1986) defines gestural compatibility in terms of the degree of mechanical constraint imposed on the articulators during production. He proposes that

> the degree of compatibility between a given gesture and adjacent gestures decreases with the degree of articulatory constraint. Thus, highly constrained gestures ought to block coarticulatory effects to a larger extent than gestures specified for lesser degrees of articulatory constraint. (p.71)

In support of this theory, he demonstrates that the degree of V-to-V coarticulation in VCV sequences is dependent upon the degree of articulatory constraint involved in both the consonantal and the vowel gestures. Recasens' speech material comprised V^1CV^2 sequences in Catalan where V= /a/ or /i/ and C= one of seven consonants involving contrasting degree of articulatory constraint on the tongue dorsum: /l, ɫ, ɾ, r, β, ð, ɣ/. Data was collected from four speakers. The greater the degree of tongue-dorsum constraint for the intervocalic consonant, the smaller the V-to-V coarticulatory effects that were observed. Effects were also found to decrease inversely with the degree of tongue-dorsum constraint for the target vowel: /i/ proved to be more resistant than /a/ to changes in tongue height and jaw opening as a function of either preceding or following vowel identity. This finding accords with the results of other studies which report greater resistance to vowel-to-vowel effects for /i/ compared with /a/ (Gay, 1974; Bell-Berti & Harris, 1976; Magen, 1989).

In addition to the amount of coarticulation, where this is defined as the extent of the displacement in target formant frequency value, Recasens (1986) also reports differences between /i/ and /a/ with respect to the timing of effects and thus the overall range of effects. Recasens found that for three of the four speakers, anticipatory effects showed an earlier onset time when V^1= /a/ than when V^1=/i/. For two speakers, carryover effects showed a later offset time when V^2= /a/ than when V^2= /i/. (Differences in the temporal extent of coarticulatory effects have also been demonstrated for lexically stressed and unstressed vowels. Magen (1989) reports that anticipatory effects begin later, and carryover effects end sooner in the case of stressed vowels.)

Recasens reports that differences between the two vowels were more apparent at the anticipatory than at the carryover level. For instance, for both /i/ and /a/, carryover effects could either last until consonantal closure or constriction or extend into V^2. However, while anticipatory effects usually started at V^1 when V^1=/a/, they did not start until consonantal closure or constriction when V^1=/i/. He therefore concludes that anticipatory effects are dependent on the degree of mechanical constraint on the tongue-body during production of the target vowel whereas carryover effects are independent of such constraints.

By extension, in an electropalatographic and acoustic study of C-to-V effects in VCV utterances, Recasens (1991) reports greater contextual variability for back vowels and for schwa than for front vowels. The difference in degree of coarticulation is similarly attributed to differences between vowels with respect to the degree of constraint on tongue dorsum activity involved during their production. Whereas, in the case of front vowels, the entire tongue body is involved in achieving a palatal constriction, for back vowels, only the back of the tongue is actively involved in the vocalic constriction, leaving the ventral surface of the tongue and much of the tongue dorsum free to coarticulate with adjacent consonants. Schwa displays the highest degree of variability of all the vowels. Accordingly, Recasens attributes this to "the inherently low requirements on the tongue configuration during the production of this vowel" (p.187).

Recasens' observations with respect to schwa are consistent with claims made by other researchers that schwa is characterised by an unstable tongue position (Kanter & West, 1960; Shriberg & Kent, 1982). In line with Keating's (1988) proposals regarding phonetic underspecification, Browman & Goldstein (1992) suggest that schwa may be completely unspecified for tongue position, that is, that it may have no independent tongue position of its own but simply represents "a specified 'interval' of time between two full vowels in which the tongue continuously moves from one vowel to another" (p. 27).

2.3.2.2 Phonetic underspecification

The recent upsurge of featural underspecification in phonology following Archangeli (1984) has resulted in a number of proposals attributing various degrees of underspecification to segments at underlying and intermediate phonological levels (Archangeli, 1988; Kiparsky, 1985; Steriade, 1987). However, while representations may be underspecified at abstract levels of the phonology, it has also been assumed that during the course of a derivation segments which lack inherent specification for given features acquire feature values through the application of phonological 'fill-in' or default rules. Thus, they are always fully specified at the surface or phonetic level. In contrast to this idea, Keating (1988) suggests that "underspecification may persist into phonetic representations" (p. 30). If a segment is phonetically unspecified for a given feature it will, in effect, be transparent to neighbouring segments' specifications for that feature. That is, rather than assimilate the feature values of neighbouring segments, it is simply "interpolated through". This idea is largely adapted from Pierrehumbert's (1980) model of intonation in which quantitative phonetic rules of interpolation build contours between target F0 values (see also Pierrehumbert & Beckman, 1988).

According to Keating, phonetic rules of interpolation produce different output forms from phonological feature-spreading rules. Transparent segments should display a more dynamic, transitional quality along the unspecified dimension that is determined equally by the preceding and following context. In contrast, segments which have acquired a feature value through phonological feature spreading should display the acquired phonetic property throughout the greater part, if not the full extent, of their duration.

Keating cites the glottal fricative /h/ as an example of a transparent segment. Being inherently unspecified for oral features, /h/ is traditionally assumed to assimilate the properties of a following vowel. According to Keating, however, "the apparent assimilatory effect on the /h/ is a dynamic, transitional one, not a static one" (p.37), the observed formant trajectories during /h/ reflecting interpolation through from the preceding to the following context segment.

Because unspecified segments make no contribution of their own to the continuous trajectory between adjacent specified segments, they also permit the interaction of adjacent context effects. Thus, in a /VhV/ sequence, it is predicted that vowel-to-vowel effects will extend in both directions through the medial /h/ such that V^1 will show systematic variation as a function of V^2 and V^2 will show systematic variation as a function of V^1.

A number of studies in the literature demonstrate strong vowel-to-vowel effects on medial schwa in VCəCV sequences, in some cases extending through the schwa (Bell-Berti & Harris, 1976; Alphonso & Baer, 1981; Fowler, 1981, 1983; Huffman, 1986; Magen, 1984, 1989). For example, Huffman (1986) found evidence of both anticipatory and coarticulatory effects during the schwa segments in /bV^1CəCV^2b/ sequences. For one speaker, carryover effects on the last full vowel from the first full vowel were also observed. Similar results were obtained by Magen (1989) who analysed trisyllabic sequences of the form /bV^1bəbV^2b/ where the flanking vowels were either /i/ or /a/. Measurements of F2 made at schwa onset, midpoint and offset showed vowel-to-vowel effects extending through the schwa in both directions. F2 values for schwa were influenced by the preceding vowel at onset and by the following vowel at offset. A high correlation was also found between variability in values measured at the midpoint of the preceding vowel and the identity of the following vowel and, similarly, between variance in the midpoint values of the following vowel and the identity of the preceding vowel.

Browman & Goldstein's (1992) study represents an attempt to determine whether observations such as these reflect interpolation through or, alternatively, may be explained in terms of the co-production of schwa with adjacent vowels. Since this is the first study which specifically addresses the question of whether or not schwa has an independent phonetic target, the methodology and results are reported in detail below.

2.3.2.3 Targeted schwa: Browman & Goldstein (1992)

Browman & Goldstein (1992) investigate the possibility that schwa is completely unspecified for tongue position in a series of stepwise multiple regression analyses using articulatory data from the Tokyo X-ray archive (Miller & Fujimura,1982) and computer simulated data. The articulatory data comprised x-ray microbeam tracks for /pV^1pə' pV^2pə/ sequences produced by an American English speaker where V^1 and V^2 represented any combination of the vowels /i, ɛ, a, ʌ, u/. Vowels were characterised in terms of displacement extrema (peaks and valleys) in the horizontal and vertical time course traces of pellets located at the tongue blade, middle and rear of the tongue-dorsum. In the case of schwa, the reference points tended to occur relatively late in its acoustic duration. Since Browman & Goldstein were interested in assessing the relative influence of V^1 and V^2 on the schwa pellet positions, they therefore also made tongue measurements earlier in the vowel at the point where lip opening for schwa was maximal.

Browman & Goldstein argue that if schwa is simply an intermediate point on a continuous tongue trajectory between V^1 and V^2, then the positions of the schwa pellets should be entirely predictable from knowledge of the pellet positions for V^1 and V^2. In order to test this, multiple regression analyses were performed on all possible subsets of three predictors: V^1 position, V^2 position and a constant factor representing an independent schwa component (the mean schwa position), to determine which (linear) combination of the three terms best predicted the position of a given pellet dimension during schwa. Results indicated that an independent schwa contribution was important to the prediction since V^2 plus the constant factor yielded a smaller residual variance for schwa than either V^2 alone or in combination with V^1. The stepwise regression also revealed an asymmetry in the amount of influence exerted by V^1 and V^2 on the tongue position during schwa. While the influence of V^1 was apparent in the initial portion of the schwa, the influence of V^2 was apparent throughout the schwa. On the basis of these findings, Browman & Goldstein posit an account of schwa's variability in terms of articulatory or "gestural" overlap. They propose that schwa does involve an active gesture but that this gesture is completely overlapped by the gesture for the following full vowel.

In order to assess the effects that would arise from varying degrees of overlap between an active gesture for schwa and the gesture for a following full vowel, Browman & Goldstein conducted a series of computer simulations using the computational gestural model developed at Haskins Laboratories (Browman & Goldstein, 1984, 1989; Saltzman et al., 1988). The model comprises three parts: (1) the gestural score which is an abstract representation of the gestures required for a given utterance and their organisation over time. This serves as input to (2) the task-dynamic model (Saltzman, 1986; Saltzman & Kelso, 1987) which produces a corresponding set of articulatory movements. These are input to (3) an articulatory synthesiser (Rubin, Baer & Mermelstein, 1981) to calculate an output speech waveform.

The simulated data was analysed in the same way as the X-ray microbeam data and subjected to the same set of regression analyses. As in the X-ray data, the best two-term prediction involved V^2 and the schwa component. Again, while there was very little variation of schwa as a function of V^1, Browman & Goldstein report considerable systematic variation of schwa as a function of V^2. V^2 contributed to the position of the pellets at both the lip and tongue reference points for schwa whereas V^1 contributed to the lip pellet position only. The presence of V^2 effects during the schwa interval support Browman & Goldstein's proposal that the schwa and the V^2 gestures are initiated simultaneously and "unfold together". The influence of V1 is attributed to inertial effects which "disappear as the tongue position is attracted to the "target" associated with the schwa and V^2 regimes" (p 51).

Browman & Goldstein performed two further sets of simulations. The gestural score for the first simulation took the same form as before with the exception that the tongue body gestures for schwa were removed: "Thus active control of V^2 began at the end of V^1, and without a schwa gesture, the tongue trajectory moved directly from V^1 to V^2. During the acoustic interval corresponding to schwa, the tongue moved along this V^1-V^2 trajectory" (p.52).

Results indicated that some warping of the trajectory was necessary to produce a schwa percept. Where V^1 and V^2 were of a different identity to each other, the resulting simulations produced utterances whose medial vowels were perceived as schwas. However, where V^1 and V^2 were of the same identity (particularly when they were high vowels) the medial vowel was perceived as being of the same quality as the full vowels.

The second set of simulations also involved an "unspecified" schwa. This time, however, the V^2 gesture was delayed so that it did not begin at the offset of the V^1 gesture but at the beginning of the bilabial closure following the schwa. Thus, during the schwa interval no tongue-body gesture was active: "The motion of the tongue-body centre during this interval, then, was determined solely by the neutral positions, relative to the jaw, associated with the tongue-body articulators, and by the motion of the jaw, which was implicated in the ongoing bilabial closure and release gestures" (p.53).

During this interval, the tongue body lowered giving rise to a schwa percept in the previously problematic /pipə 'pipə/ sequence. However, lowering was also apparent in the other utterances including /papə ' papə/ which, in the X-ray data, witnessed a raising of the tongue-body during the schwa interval. Browman & Goldstein argue that this shows that the neutral positions for individual articulators do not produce the same articulatory effect as an active schwa gesture towards an overall neutral configuration. They interpret this as further support for their claim that schwa is specified for tongue position. Overall, they conclude that schwa does have a target but that "it is still "colourless", in that its target is the mean of all the vowels, and is completely overlapped by the following vowel" (p.54).

2.3.2.4 Targetless schwa: Van Bergem (1994)

In a more recent acoustic study of coarticulatory effects on schwa, Van Bergem (1994) presents evidence in support of the targetless analysis. He examines Dutch schwa in ˈVCəC and CəˈCV nonsense words where the context vowel is one of /i, a:, u/ and the consonants any two of /p, t, k, f, s, χ, m, n, ɳ, r, l, j, ʋ/. Data was collected from three speakers and comprised 897 test words in total. Van Bergem reports highly systematic coarticulatory effects on the schwa F2 value which could be described using a simple linear model. 82% of the variance in schwa second formant trajectories was explained with second-order polynomials, using the vowel onset, midpoint and offset as the three reference points. A good approximation of the formant trajectories was also obtained with first-order polynomials using just onset and offset values. In three-way analyses of variance, between 29% and 38% of the variance in schwa F1 value and between 72% and 79% of the variance in schwa F2 value was explained by context. On the basis of these results, he concludes that schwa is "a vowel without articulatory target that is completely assimilated with its phonemic context" (p.143).

The conflicting conclusions presented in Browman & Goldstein's (1992) and Van Bergem's (1994) study may reflect differences between the two studies with respect to speech material and methodology. However, Browman & Goldstein's argument has also been weakened through criticism of their evaluation of the regression results (see Kingston, 1992, same volume) and of the distinction they make between an overall neutral rest configuration and a neutral configuration specific to schwa (Barry, 1992, same volume). Both points of criticism are discussed more fully in section 5.4. In the absence of comparative data for the full vowels, the proportion of explained variance Van Bergem cites as evidence of schwa's targetlessness is a fairly arbitrary measure. It is, therefore, also possible that the high degree of context-dependency he observes for schwa reflects weak specification as opposed to complete underspecification for tongue position.

2.3.2.5 Window model of coarticulation

The possibility that schwa may be weakly specified as opposed to completely unspecified for tongue position, is accommodated within Keating's (1990) window model of coarticulation which represents the development of her earlier ideas concerning phonetic underspecification. The major advantage of the new model is that it allows for intermediate degrees of specification and is therefore no longer an "all-or-nothing" option. It thus also permits an account of vowel reduction in terms of phonetic underspecification. Whereas an all-or-nothing approach would imply a discrete alternation between schwa and the full vowels which, in turn, implies articulatory and acoustic centralisation, the window model framework accommodates inherent differences between vowels with respect to the degree of contextual variability they exhibit.

In the window model, speech production is no longer viewed as a process of connecting up target values. Instead, segments are characterised by "windows" which represent the full range of contextual variability they exhibit along given articulatory dimensions. During production, rules of interpolation serve to build paths through the windows.

Windows are determined empirically by finding the maximum and minimum value possible for a given dimension across all contexts. Window width is thus directly proportional to overall degree of contextual variability and subsequently varies according to individual segments. Segments which display a high degree of variability are characterised by a wide window and segments which display little variability are characterised by narrow windows.

Windows are not subject to further variation as a function of context since they represent the full range of variability possible for a given feature. Variation across utterances occurs as a function of the different combinations of wide and narrow windows. Depending on the sequence of windows, the path through any given window may span the entire window width or a more limited range of values within the window. Both window width and the relative position of the windows within the specified dimension serve to determine the path through them. This is illustrated in Figure 2.1, taken from Keating (1990), which shows a sequence of windows within a single articulatory dimension and the schematic interpolation between them.

In (a), a segment characterised by a narrow window occurs between two segments of the same identity to each other which are characterised by wider windows. In this case, the middle segment shows little variation across contexts and also exerts most influence on the interpolated trajectory. In (b), the middle segment is characterised by a wider window than the two context segments. The middle segment assimilates its turning point to the context and in some cases may simply be interpolated through. In (c) and (d), wide and narrow windows occur in asymmetrical contexts. Whereas the wide window shows a linear interpolation between many different contexts, the narrow window serves to constrain the interpolation.

Keating formulates the window model principles for a single articulatory dimension. However, she claims that the same principles could also be applied to acoustic parameters. Although much of the fine working detail has still to be developed, the window model, potentially, provides a framework for describing gradient coarticulatory phenomena.

2.4 Acoustic and perceptual constraints on variability

Hypotheses regarding the inherent variability of vowels have also been made on the basis of acoustic and perceptual considerations (Stevens, 1972, 1989; Syrdal & Gopal, 1986; Bladon et al., 1984; Lindblom, 1983, 1990; Lindblom et al., 1992). Stevens (1972; 1989) proposes that there are areas in the vocal tract which show quantal relations between articulation and acoustic output, that is, where perturbations in articulation produce minimal effects on the acoustic output. These regions correspond to the regions which yield closely spaced formant frequencies and hence well-defined spectral peaks (Fant, 1956). Low variability is therefore predicted for /i/ on account of the close proximity between F2 and F3 and for /ɑ/ and /u/ on account of the close proximity between F1 and F2, /ɑ/ being characterised by high F1 and low F2 target values and /u/, by low target values for both F1 and F2. The quantal theory of speech production is discussed in more detail in Chapter 7.

Syrdal & Gopal (1986) also predict lower variances for vowels with closely spaced formants than for vowels with widely separated formants. Their predictions, however, are made on the basis of perceptual constraints on variability. According to the bark-difference model, feature categories are defined by a critical distance between formants and between F1 and F0 of 3 Bark. Front vowels are characterised by F3-F2, and high and mid vowels by F1-F0, bark-difference values of within 3 Bark. Back vowels and low vowels have F3-F2 and F1-F0 bark-difference values in excess of 3 Bark. In order that the critical distance metric be met, "proximal formants may vary only within a 3-bark range from one another, but more distant formants have a much wider range over which to vary while still maintaining a distance greater than 3 Bark" (p.1095). This idea is explored further in Chapter 4.

Syrdal & Gopal's prediction is the opposite of that predicted by Bladon et al., (1984). According to Bladon et al., vowel recognition is achieved by a process of auditory spectral template matching. Since formant spacing is an important part of the matching procedure, greater variability is tolerated for proximal, integrated formants than for relatively isolated formants.

Lindblom's theory of Hyper-hypo speech production (Lindblom, 1983, 1990; Lindblom et al., 1992) also makes predictions regarding vowel variability on the basis of perceptual considerations. However, in contrast to Stevens' (1972, 1989) quantal theory or the proposals of Syrdal & Gopal (1986) and Bladon et al., (1984), the Hyper-hypo theory of speech production seeks to account for style-conditioned variability rather than inherent differences between vowels. The Hyper-hypo theory of speech production is discussed in more detail in section 7.3.

2.5 Phonological vowel reduction

In view of the fact that many diachronic patterns of variation represent fossilisations or phonologisations of synchronic coarticulatory patterns of variation (Ohala, 1993), the relative magnitude of coarticulatory effects for different vowels may also be predictable on the basis of the extent to which they alternate with schwa at the morpho-phonemic level.

The morpho-phonemic alternation of full vowels with schwa and/or /ɪ/, referred to here as phonological (as opposed to phonetic) vowel reduction, may be either obligatory or non-obligatory. Obligatory phonological vowel reduction refers to the process by which, within the generative tradition, an underlying full vowel reduces to surface [ə] and/or [ɪ]. The alternation between [e] and [ə] in the second syllable of "explain" [ɛksplen] and "explanation" [ɛkspləneʃən] is an example of obligatory phonological vowel reduction. Similarly, in derivational pairs such as the following: "photo<u>graph</u>ic, photo<u>graphy</u>"; "diplo<u>mat</u>ic, diplo<u>macy</u>"; "tele<u>graph</u>ic, tele<u>graphy</u>", a full vowel appears in the stressed syllables and schwa in the unstressed syllable. Standard generative phonology posits a unique morphological base form from which each of these related forms are derived via an ordered set of stress rules. (For a full exposition of these, see Chomsky & Halle, 1968.) A similar treatment is given in the more recent metrical phonology (Liberman & Prince, 1977; Hayes, 1984).

In contrast to phonetic vowel reduction, obligatory phonological vowel reduction is independent of local context, stress and rate effects. It is a "stable" part of the phonology (Bolinger, 1985), occurring whenever the appropriate morpho-phonological conditions are met. The standard or 'official' pronunciation of "explanation", "photography", "diplomacy" and "telegraphy" is with /ə/ in the unstressed syllables. Thus, reduction in vowel quality is obligatory insofar as the forms [ɛkspleˈneʃən], [foʊˈtɒgrafɪ], [dɪˈploʊmasɪ] and [tɛˈlɛgrafɪ] are unacceptable.

Non-obligatory phonological vowel reduction concerns the surface alternation of full vowels with [ə] and/or [ɪ]. For example, the following words may be pronounced with either a full vowel or schwa in the unstressed syllable: "molecular" ([moʊˈlɛkjulə] versus [məˈlɛkjələ]), "condensation" ([kɒndɛnˈseʃən] versus [kɒndənˈseʃən]), "september" ([sɛpˈtɛmbə] versus [səpˈtɛmbə]), "eccentric" [ɛkˈsɛntrɪk] versus [ɪkˈsɛntrɪk]), "anecdote" ([ˈanɛkdoʊt] versus [ˈanɪkdoʊt]). Other examples include the alternation between [ɪ] and schwa in "derision", "ferocity", "believe", "tranquil"; between [ɒ] and [ə] in word-initial syllables such as "botanic", "morality" and "collate"; between [ɜ] and [ə] in "conurbation", "interdict", "verbose"; between [a] and [ə] in "abdominal", "acceleration", between [ʌ] and [ə] in "conductivity", "consultation", and between [ʊ] and [ə] in "beautiful" and "literature".

In contrast to obligatory phonological reduction, non-obligatory reduction is conditioned by factors such as speech rate and style, reduced forms tending to be more common at faster speaking rates or in more casual speaking styles (Dalby, 1984). Reduction also tends to occur more frequently in function words and grammatical items which carry relatively little semantic weight (Bolinger, 1975). Thus the widespread alternation of /ɪ/ with schwa in unstressed position may be largely attributed to the relatively high proportion of suffixes in which /ɪ/ occurs: for example, /ɪbəl/ "indelible, intelligible, possible", /ɪlɪ/ "gloomily, heavily, merrily" and /ɪtɪ/ "equality, morality, legality" as well as the plural suffix /-ɪz/, the /-ɪŋ/ suffix indicating the present participle and the /-ɪd/ suffix indicating the past tense.

Non-obligatory phonological vowel reduction also operates along a tense-lax continuum such that while lax vowels alternate with [ɪ] and/or [ə] in unstressed syllables, tense vowels tend to alternate with their lax cognates, although further reduction to schwa may be possible in some cases at faster speaking rates. For example, [i] alternates with [ɪ] in "demobilise", "precursor", "eclectic", "eject", "Hebraic", "legality"; [u] alternates with [ʊ] in "euphoria", "musician", "ubiquitous", "capsule" and "fortune" and [ɔ] alternates with [ɒ] in "alternate", "already", "also", "alter", "cobolt", "salt". Of all the vowels, the low, back, tense vowel /ɑ/ shows the least tendency to reduce: "arcade", "arboretum", "demarcate", "embarkation", "rampart", "impart".

The existence of alternative lexical variants and their apparent style conditioning is indicative of sound change in progress (Ohala, 1981, 1989). It is known that connected speech processes may change over time from being gradual, continuous effects to being discrete 'phonologised' rules (see Kerswill, 1985). The fact that this change takes place gradually (Linell, 1979) allows for the possibility of processes existing at different time points along the continuum. That is, at a given time, the 'old' and 'new' forms of the units undergoing change will both be in current usage within the speech community. The selection of one variant over another is likely to be conditioned either by sociological differences between speakers or by stylistic considerations within the speech of an individual or group of speakers.

With respect to vowel reduction, a vowel that is subject to a high degree of phonetic reduction in a given word may be perceived as schwa. As Van Bergem (1995) demonstrates, listeners are often not able to unambiguously classify vowels as either full or reduced, particularly if they occur in an unstressed syllable that is intermediate between a primary and secondary stressed syllable. Over time, this perceptual ambiguity may lead to the existence of two lexical variants, one with a full vowel and one with a schwa. The form with a schwa may be more commonly used in casual speech and the full vowel variant in more formal speech modes. Eventually, the variant form with schwa may become the only acceptable form. Thus, obligatory and non-obligatory phonological vowel reduction and phonetic vowel reduction might be viewed as part of a single historical continuum of reduction.

According to Van Bergem (1990), phonological and phonetic vowel reduction (in his terms, 'lexical' and 'acoustic' vowel reduction) represent a means of economising on articulatory effort:

> ... people tend to relax their articulation in less informative parts of an utterance which results in vowels with a relatively short duration and a relatively large spectral distance from the 'ideal vowel target'. The spectral shift is usually toward the schwa position, indicating that the schwa probably

> requires the least amount of articulation in many consonantal contexts ... It is only a small step to replace such a vowel that has hardly any identity with a schwa and to make this substitution a permanent part of the lexical system. (p. 10-3)

Accepting this relationship between phonetic and phonological reduction, it should be possible to observe parallels between the two 'processes' with respect to the pattern of variability across vowels. For example, the operation of non-obligatory phonological vowel reduction along a tense-lax continuum implies that lax vowels are more "cost-effective" than tense vowels which, in turn, implies that they are less narrowly specified. Assuming this is correct, lax vowels might be expected to display greater phonetic variability generally than tense vowels (i.e. in both stressed and unstressed position). Comparable levels of variability to schwa might also be predicted for /ɪ/ in view of its extensive alternation with schwa and its own 'reduced' status (Fudge, 1984).

2.6 Summary and research objectives

In this chapter, alternative accounts for observed differences in the degree of phonetic variability displayed by different vowels are explored. Higher levels of contextual variability noted in the literature for schwa compared with the full vowels and for lax compared with tense vowels suggest a hierarchy of vowel robustness in which tense vowels show the greatest, and schwa the least, resistance to coarticulatory effects.

Adopting the traditional target undershoot approach to coarticulation, greater contextual variability for lax vowels than for tense vowels may be attributed to the shorter inherent duration of these vowels. However, within a gestural framework, it is also possible that lax vowels are more susceptible to coarticulatory effects because they are inherently less fully specified than their tense counterparts where this implies less articulatory effort and/or less divergence from neutral. In this case, shorter durations may be concomitant with reduced vowel quality but are not the chief determinant of the reduction.

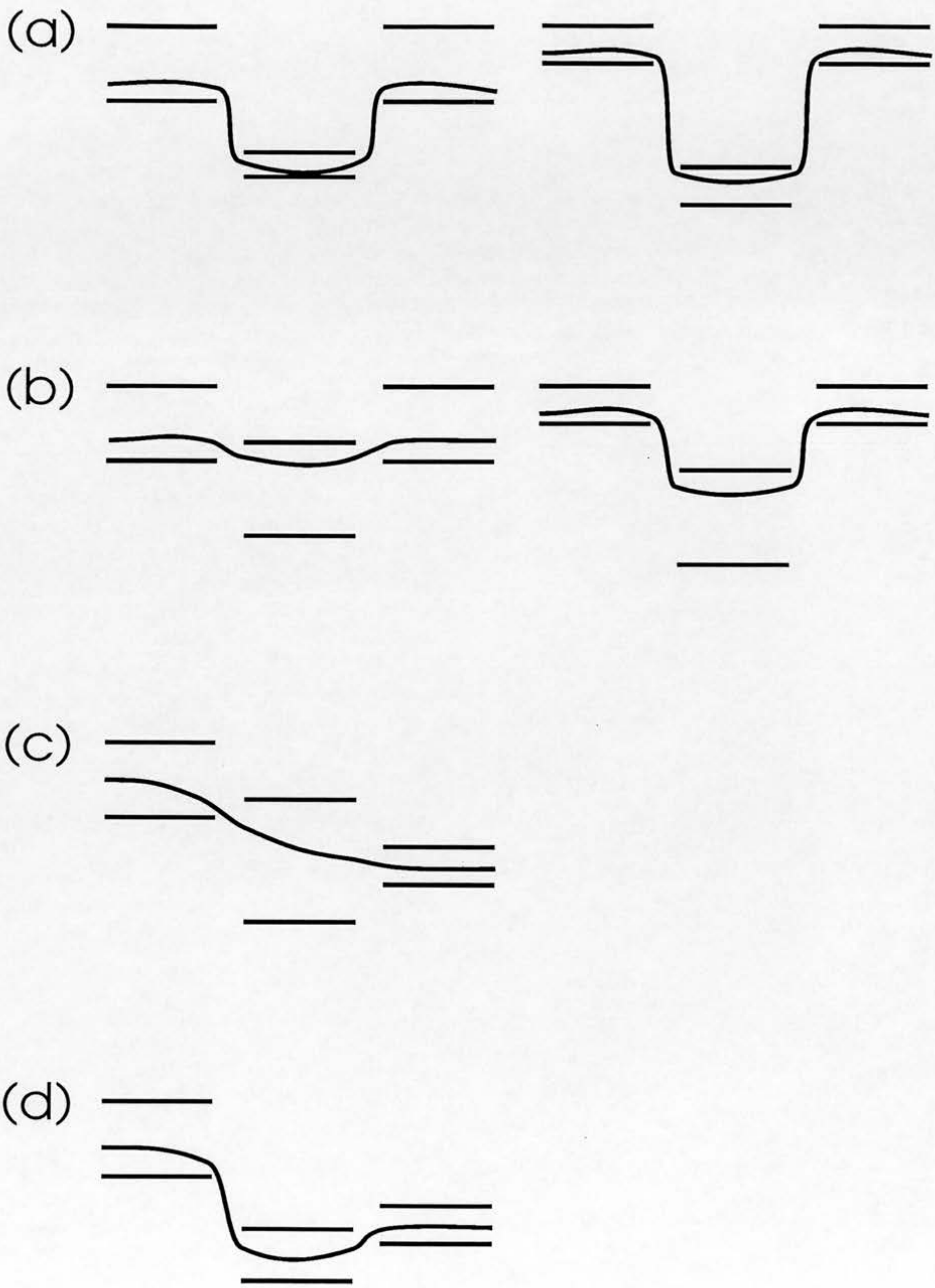
It is not the intention of this thesis to explicitly test Lindblom's (1963) Undershoot Hypothesis. This would require a controlled experimental analysis of the data which is not within the scope of the present study. Furthermore, there already exist a large number of studies in the literature which show that vowel quality may vary independently of duration (Verbrugge & Shankweiler, 1977; Gay, 1978; Tuller, Harris & Kelso, 1982; Lisker, 1984; Nord, 1986; Summers, 1987; Den Os, 1988; Engstrand, 1988; Fourakis, 1991; Van Son, 1993 inter alia.). Lindblom himself has also amended the original model to accommodate variation in the degree of duration dependent undershoot as a function of speaking effort and style (Lindblom, 1983; Lindblom & Moon, 1988, Lindblom, 1990; Lindblom et al., 1992). Thus, accepting that differences between vowels (full versus reduced, stressed versus unstressed, tense versus lax) cannot be accounted for solely in terms of durational differences, this thesis aims to establish whether there is empirical support for an account of vowel reduction in terms of phonetic underspecification.

In the following analysis (see Chapters 5 and 6), vowels are compared with respect to the magnitude and patterns of contextual variability they exhibit with reference to the following three questions:

(1) What evidence is there to support the suggestion that schwa is completely unspecified for tongue and jaw position?
(2) Are unstressed vowels more sensitive to context effects than their stressed counterparts?
(3) What evidence is there to support a tense/lax, front/back or high/low distinction between vowels in terms of inherent variability?

Evidence that schwa is targetless and completely dependent on context would support the view of vowel reduction as a means of economising on articulatory effort since the straight-line interpolation between consonants arguably reflects minimal articulatory effort. Greater contextual assimilation for unstressed vowels compared with stressed vowels and for lax vowels compared with tense vowels might therefore also be interpreted as a saving on articulatory effort rather than as an automatic consequence of shorter durations. A front/back distinction between vowels in terms of variability (Recasens, 1991) may also be explained in terms of varying degrees of specification and contextual assimilation. It is, however, less consistent with the view that phonetic and phonological vowel reduction in English are part of the same historical continuum.

Figure 2.1: A sequence of windows and the schematic interpolation between them.



Chapter 3
Description of the Data-base

3.1 Corpus

The speech material used in this thesis comprises six hundred and sixty sentences read by one Southern British Standard (SBS) male speaker. These represent a combination of two hundred sentences from the ATR speech data-base (Laver et al., 1989) and a further four hundred and sixty sentences from an anglicised version of the TIMIT speech data-base (Lamel et al., 1986).

The sentences cover almost the full range of permissable syllables in English, where the syllable is defined as an obligatory vowel nucleus optionally preceded by between one and three consonants and optionally followed by between one and four consonants. They also give at least one example of all phonotactically permissible sequences of initial and final consonants and at least one example of all possible consonant-vowel and vowel-consonant combinations. For a full description of the design considerations and a complete listing of the sentences, see Laver et al. (1988). Ten example sentences (five ATR and five TIMIT) are given below:

(1) The price range is smaller than any of us expected.
(2) The smell of the freshly ground coffee never fails to entice me into the shop.
(3) Thank goodness it's Friday and time to go home.
(4) It's difficult to choose between two such equally good alternatives.
(5) He thanked his colleague for the invaluable lifts to the various conferences that term.

(6) The big dog loved to chew on the old rag doll.

(7) Withdraw only as much money as you need.

(8) My desires are simple: give me one informative paragraph on the subject.

(9) The speech symposium might begin on Monday.

(10) The surplus shoes were sold at a discount price.

3.2 Recording and digital data acquisition

The sentences were stereo recorded at the Centre for Speech Technology Research, Edinburgh (CSTR) on a Sony PCM-F1 Digital Audio Unit in a sound treated room using a Shure SM-10 A close-talking dynamic cardioid microphone and a desk-mounted Sennheiser MKH 406 P48 condenser cardioid microphone. The recordings were low-pass filtered at 10 kHz and digitised at a 20 kHz sampling frequency with 12 bit resolution.

The same procedures were used for both the ATR and TIMIT data. Although there was a time-lapse between recording sessions for the two sets of sentences, comparison of vowel formant frequencies in both sets, showed that there was no significant difference in vowel formant frequency value as a function of set membership.

3.3 Segmentation

The sentences were hand-segmented and labelled in a broad phonetic transcription by a number of phonetically trained transcribers (including the author) according to a fixed set of segmentation and labelling criteria (see Laver et al., 1989). This was done using CSTR's Audlab system incorporating information from a spectrogram, a time-amplitude waveform and auditory feedback. While it is impossible to guarantee 100% segmentation accuracy (Peterson & Lehiste, 1960; Umeda, 1975; Klatt, 1975), all transcriptions were proof-read and amended where necessary by the author in order to ensure as high a level of consistency as possible. Vowel tokens were also independently hand-labelled for sentence stress by a single transcriber (see section 6.1).

3.4 Formant estimation and tracking

Formant values were obtained using a peak-picking algorithm which employs a global optimisation based on a generalisation of the centroid (Crowe, 1987). According to this method, the speech material is Fourier-transformed using a 512-point FFT with a 20 ms Hamming window, moved along at 5ms intervals. "High frequency emphasis (6dB/octave) is applied to the power spectrum, and the triple centroid is computed over the frequency band from 250 Hz to 2875 Hz" (p. 1019). Formant tracks are produced by connecting the frequency estimates with no smoothing or post-processing (see Crowe, 1987 for a full description). F0 values were obtained using a pitch tracking algorithim developed at CSTR (Bagshaw et al., 1993).

3.5 Data-processing

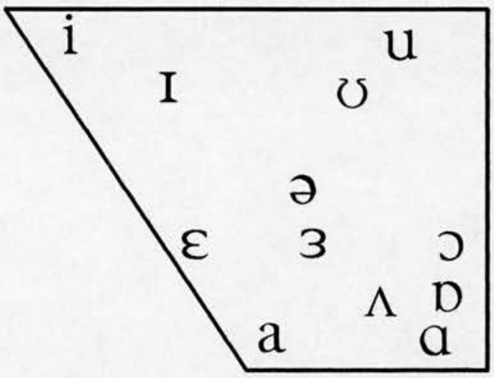
The data was processed using Acoustic Phonetic S (APS). APS was developed at CSTR under the Alvey Speech Input Project (Thomson & Laver, 1987) and consists of a set of extensions to the S interactive data-analytic and graphics system (Becker & Chambers, 1984). APS includes sampling functions which return the mean, maximum or minimum values for a specified acoustic interval as well as various plotting and statistical functions. As an example of the type of data manipulation that is possible with this system, APS commands can be used to access all the /ə/ segments in the data-base together with their left and right contexts, sample the formant values at the vowel midpoint and/or at any other specified time point and compute their durations. For a full description of APS see Watson (1988).

The majority of the figures in this thesis were created using the Microsoft Excel data-analysis and graphics package. The Systat statistical package was also used for the statistics in Chapter 4. In all other cases, the BMDP statistical software was used. A full description of the statistical procedures are given in the appropriate analysis sections.

3.6 Accent system

The Southern British Standard accent system comprises twelve monophthongs /i, ɪ, ɛ, a, ə, ɜ, ɑ, ʌ, ɒ, ɔ, ʊ, u/ and eight diphthongs /eɪ, aɪ, aʊ, oɪ, oʊ, iə, ɛə, uə/. In the present study, only the monophthongs are used in the analysis. In Figure 3.1, the position each of these occupies on the traditional articulatory/acoustic vowel quadrilateral is shown.

Figure 3.1: Vowel quadrilateral for the SBS accent system



This thesis follows convention in using measures of the first and second formant frequencies to characterise vowels. While it is a matter of debate as to whether formant frequencies are the optimal descriptors of vowel quality (Bladon, 1982), the correspondences between formant frequency value and the traditional phonological and phonetic dimensions of height and front/backness are well established (Stevens & House, 1956; Fant, 1960; Veatch, 1991; Maeda & Honda, 1994). F1 is negatively correlated with vowel height such that low vowels are characterised by relatively high F1 values while high vowels are characterised by relatively low F1 values. Vowel front/backness is related to F2 frequency, front vowels being characterised by relatively high F2 values and back vowels being characterised by relatively low F2 values. The frequency value of F2 also reflects the lip-rounding feature, with lower frequencies accompanying increased lip-rounding.

In Table 3.1, mean formant values for vowels in the present data are listed for comparison with data from Wells (1962). Mean formant values reported by Peterson & Barney (1952) are also included for comparison. The present data was sampled at the durational vowel midpoint (see Chapter 4 for a discussion of sampling procedure) and, for each vowel, represents the mean values for tokens pooled across all contexts. Both Wells' and Peterson & Barney's data represent mean values for isolated productions of vowels in /hVd/ sequences. Peterson & Barney's data is averaged across a sample of thirty three American English male speakers, while Wells' data is averaged across a sample of twenty five British English male speakers. The values from Wells' and the present study are also plotted in Figure 3.2.

Table 3.1: A comparison of mean first and second vowel formant values (Hz) for SBS and American English vowels. An asterisk marks those cases where values are not available.

	Present data		*Wells '62*		*P&B '52*	
	F1	*F2*	*F1*	*F2*	*F1*	*F2*
i	372	2080	300	2300	270	2290
ɪ	442	1722	360	2100	390	1990
ɛ	607	1633	570	1970	530	1840
a	756	1503	750	1750	660	1720
ə	481	1445	*	*	*	*
ɜ	555	1380	580	1380	*	*
ɑ	698	1108	680	1100	730	1090
ʌ	663	1195	720	1240	640	1190
ɒ	598	1014	600	900	*	*
ɔ	469	845	450	740	570	840
ʊ	428	1326	380	950	440	1020
u	371	1570	300	940	300	870

The generally lower F2 values for front vowels and higher F2 values for back vowels in the present study may be attributed to the global centralising effects of adjacent segmental context. Higher F1 values for the high vowels may similarly reflect contextual influence. Despite differences in absolute frequency value, which are to be expected across different speakers and contexts, the same relationship between vowel categories with respect to height and front/backness is evident in each data-set with the exception of the two high, back vowels /ʊ/ and /u/. These vowels are realised by the present speaker with a more fronted quality and are therefore characterised by generally higher F2 values than in Wells' or Peterson & Barney's data. However, Wells (1962) also observes that /u/ and /ʊ/ are sometimes centralised. Henton (1983) also notes that the fronting of these two vowels is becoming increasingly common among younger speakers of SBS. In the following analysis /u/ and /ʊ/ are considered separately in front-back comparisons of the data.

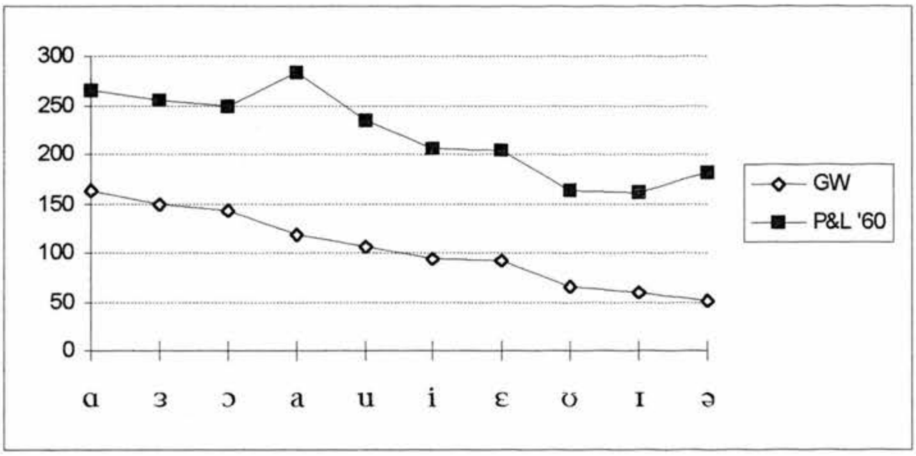
Mean durations for vowels in the present data are given in Table 3.2. They are also represented graphically in Figure 3.3 together with the durations reported by Peterson & Lehiste (1960) for comparison. ('GW' represents the present speaker's initials.)

Table 3.2: Mean durations for vowels pooled across all contexts (msec).

	ɑ	ɜ	ɔ	a	u	i	ɒ	ɛ	ʌ	ʊ	ɪ	ə
Mean	164	148	143	118	106	93	100	92	89	66	59	51
SD	47	40	46	34	47	34	35	31	28	23	22	24

The large difference in absolute value between the mean durations for the vowel data in this study and the mean durations reported by Peterson & Lehiste (1960) reflect differences in speech material. Peterson & Lehiste examine vowels in isolated CVC syllables produced in the same carrier phrase. Durations for vowels in connected speech are influenced by a wide range of factors including segmental and suprasegmental context (Klatt, 1973, 1976) and are typically much shorter in duration than vowels produced in isolation. Despite differences in absolute value, the two studies generally show the same hierarchy of inherent vowel duration. In both cases, longer mean durations are obtained for the low and mid-low, tense vowels /ɑ, ɔ, ɜ/ and the long, lax vowel /a/ than for the high, tense vowels /i, u/ which, in turn, show longer mean durations than the lax vowels /ɪ, ɛ, ʊ, ə/. In the present data, the lax vowel /ʌ/ patterns with the other lax vowels while /ɒ/ shows a similar mean duration to /i/ and /u/.

Figure 3.3: A comparison of mean durations for present vowels (GW) and data reported by Peterson & Lehiste (1960).

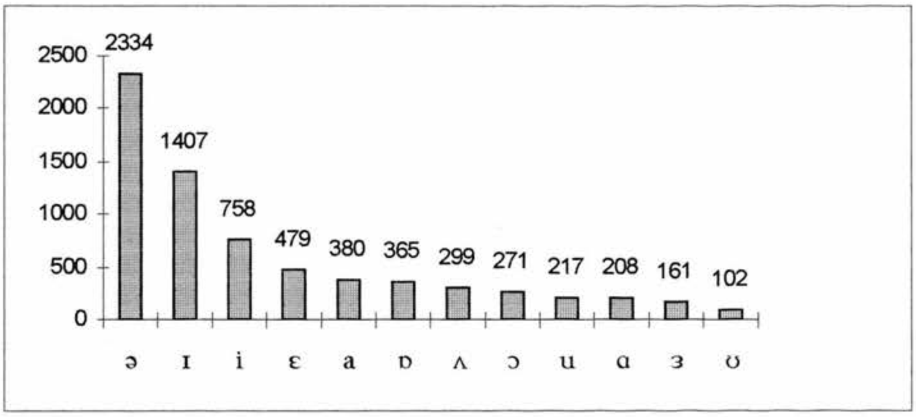


A one-way analysis of variance performed for the present vowel data showed a significant main effect of vowel identity for both the set of long vowels ($F(5, 1195) = 142.85$, $p < .0001$) and the set of short vowels (excluding schwa) ($F(4, 2653) = 270.68$, $p < .0001$). Post hoc Tukey pairwise multiple comparison of means tests showed no significant difference in duration between the long vowels /ɔ/ and /ɜ/ or between the short vowels /ɛ/ and /ʌ/. Otherwise all pairwise differences within each set were significant at either $p < .01$ or $p < .001$. A comparison of the vowels /a, i, u, ɛ, ʌ, ɒ/ showed no significant difference in duration between /u/ and /ɒ/ or between /i/, /ɛ/ and /ʌ/. The difference in duration between /i/ and /ɒ/ was significant at $p < .05$. All other pairwise comparisons were significant at either $p < .01$ or $p < .001$.

3.7 Distribution of phonemes

The total number of vowel tokens used in the analysis is 8579. Figure 3.4 shows how these are distributed across vowel categories.

Figure 3.4: Distribution of vowels

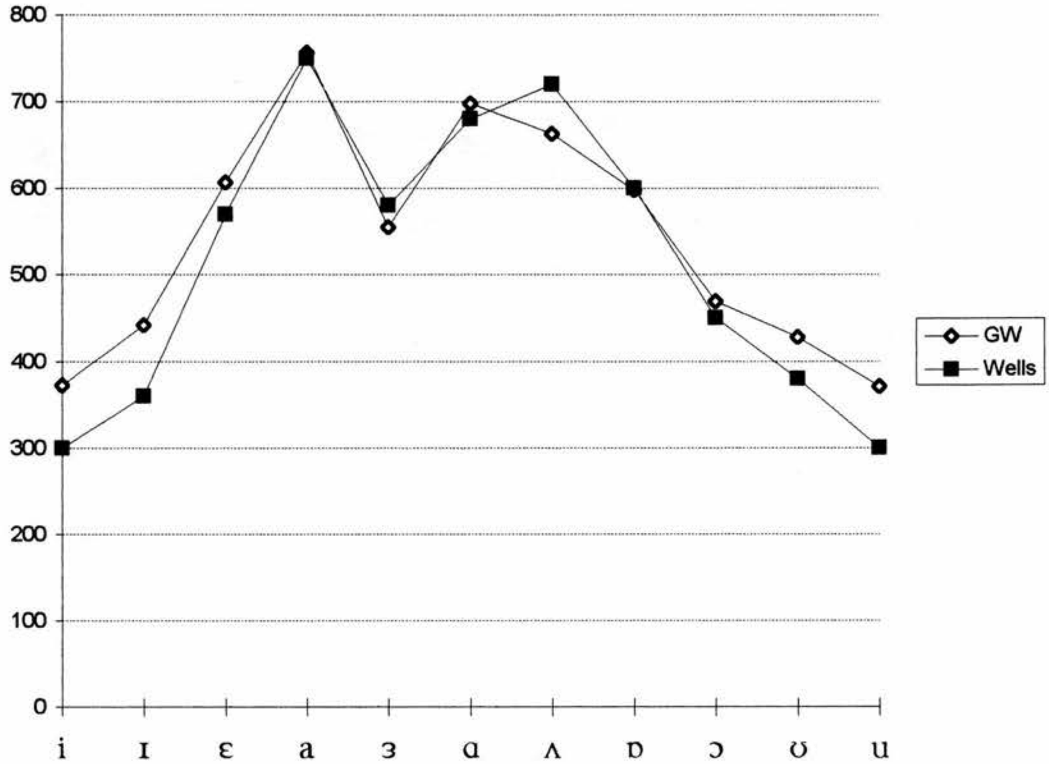


Preceding and following context segments for each vowel are listed in Tables 3.3 and 3.4. Individual contexts expressed as a proportion of the total number of contexts per vowel are also represented graphically in Figure A 1, Appendix A.

The choice of sampling point and means of representing formant frequency values in this study are discussed in detail in the following chapter. Details of other measurement procedures are given in the appropriate analysis sections.

Figure 3.2: Comparison of mean formant values for SBS monophthongs.
GW refers to the present speaker and Wells to data from Wells (1962)

3.2a: **F1**



3.2b: **F2**

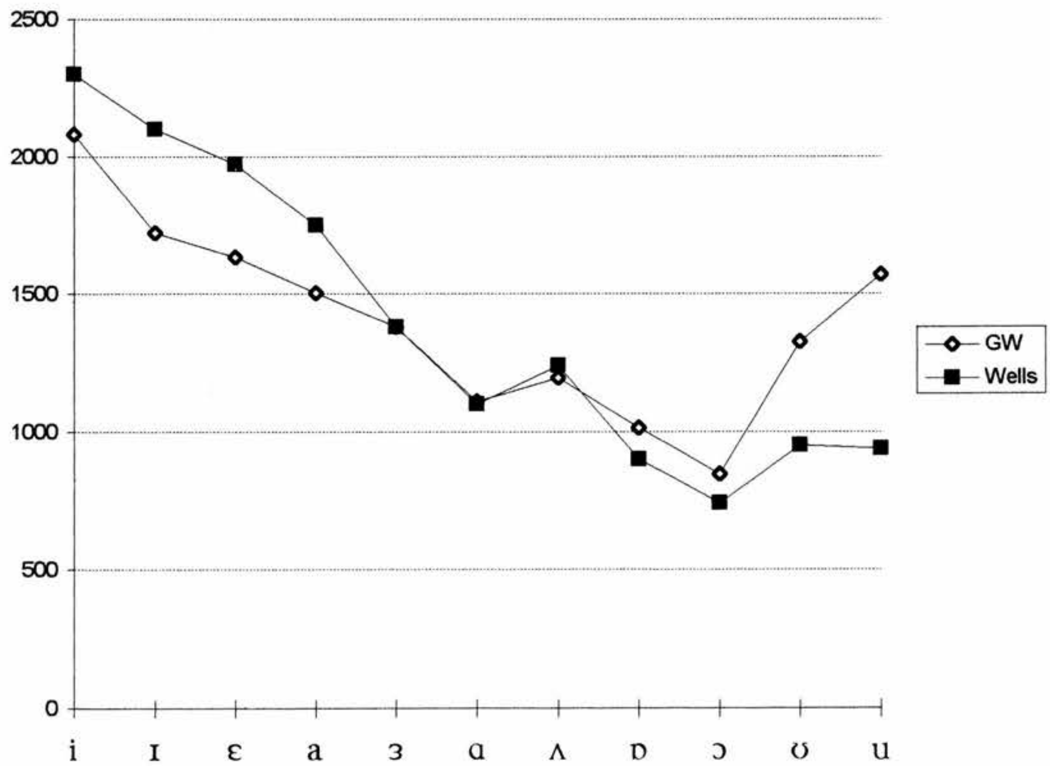


Table 3.3: Distribution of preceding contexts

Context	i	ɪ	ɛ	a	ə	ɜ	ɑ	ʌ	ɒ	ɔ	ʊ	u
s z	51	188	37	16	201	23	10	26	19	17	1	16
n	55	74	32	12	105	6	4	19	31	9	1	2
t d	90	293	47	36	404	17	17	22	48	25	15	34
l r	257	228	95	70	221	7	40	51	75	22	11	36
ð θ	65	56	6	16	449	6	1	6	1	4	0	1
m	23	51	24	25	74	5	4	38	11	22	0	8
p b	63	81	34	35	115	24	27	17	38	22	13	6
f v	29	72	28	18	149	19	12	8	13	35	7	4
w	31	103	51	3	98	16	2	18	23	30	7	3
ʃ ʒ	31	65	32	12	114	10	11	12	16	13	9	23
j	0	1	9	2	47	1	3	7	5	2	19	68
ŋ	1	22	1	1	8	1	1	0	4	1	0	0
k g	15	50	31	54	126	8	34	55	35	31	13	11
Vowel	16	80	30	31	172	7	17	10	31	26	1	1
ʔ h	24	20	15	36	17	10	12	9	11	8	5	4
Silence	7	29	7	13	34	1	13	1	4	4	0	0

Table 3.4: Distribution of following contexts

Context	i	ɪ	ɛ	a	ə	ɜ	ɑ	ʌ	ɒ	ɔ	ʊ	u
s z	95	293	46	16	335	21	24	24	29	30	3	32
n	40	196	101	63	382	15	20	65	70	16	1	14
t d	110	208	72	63	306	39	55	31	65	67	29	34
l r	37	76	91	53	283	9	15	31	53	60	19	11
ð θ	14	53	7	5	56	11	7	20	2	12	1	15
m	24	58	21	38	146	7	16	32	12	14	4	8
p b	57	40	13	31	228	15	12	32	52	7	2	12
f v	46	74	44	21	204	16	29	24	27	5	0	11
w	24	7	0	0	75	3	2	0	0	12	0	5
ʃ ʒ	9	60	15	14	21	2	9	5	6	2	4	5
j	3	0	0	0	16	0	1	0	0	4	0	2
ŋ	5	158	5	16	1	0	0	20	14	1	0	0
k g	55	170	62	56	168	14	16	15	33	16	23	10
Vowel	178	14	1	1	28	7	0	0	0	19	15	53
ʔ h	11	2	0	0	45	0	1	0	1	3	0	3
Silence	52	4	0	0	40	2	1	0	1	3	1	2

Chapter 4
Representation of the Spectral Vowel Target

Introduction

In this chapter, alternative single-number representations of vowel formant values commonly employed in the literature to represent the spectral vowel target are compared. The aim is to establish which of these measures is the most suitable for representing the vowels in the present data.

It is debatable whether vowels may be satisfactorily characterised by a single spectral cross section (Strange, 1989) or indeed whether formant frequencies are the optimal descriptive parameters of vowel quality (Bladon et al., 1982). However, a single number representation of the formant trajectory is considered desirable for present purposes on two counts; firstly to facilitate comparison of results with findings in the literature (the majority of coarticulatory studies to date describe contextual effects in terms of shifts in vowel formant frequency measured at a single instant in time), secondly, because it is more convenient and computationally less expensive than other multi-parametric or dynamic representations. The decision to use a single number representation was considered justifiable on these grounds given that what is required for this thesis is a suitable comparative measure of vowel quality, not a complete description of the vowel target per se.

The survey is motivated by the lack of consensus in the literature with regard to the question of where the formants should be measured and how the frequency values should be represented. Alternative sampling points include the durational vowel midpoint, the vowel steady-state and the turning point in the formant trajectory where this represents movement toward and away from a maximum point of displacement. Frequency values may be represented linearly or non-linearly along a scale such as the Mel or Bark scale, in terms of spectral distance measures or as absolute frequencies. It has yet to be established which of the alternative methods of representation provides the best classification. The question of whether or not different methods are biased towards different vowels also remains to be investigated.

A further question concerns the practical applicability of different methods of representation. Since empirical studies of vowel quality to date almost exclusively use speech material which is carefully controlled with respect to context and stress, it is not yet known how far the descriptive techniques employed in these studies can be applied to vowel data which is unrestricted with regard to these variables. In natural continuous speech, vowels are typically shorter in duration and spectrally more variable than when pronounced in isolation or in a careful laboratory style in which case the location of steady-states may not always be possible. Trajectory shape is also likely to vary depending on vowel duration, on the nature of the surrounding context or on a combination of both these factors.

The extent to which methods widely used in the literature can be applied to a sample of vowel data where context and stress is not carefully controlled is clearly an important issue with respect to the purposes of the present study. It also has important implications with regard to the generalisability of coarticulatory theories that have been developed on the basis of observations of preselected vowel data.

The first part of the chapter addresses the question of where the formants should be measured. Alternative sampling points are compared with respect to their performance in classifying the vowels in the present data. Success in characterising vowels is measured in terms of the statistical separability of vowel distributions (Assman et al., 1982; Syrdal & Gopal, 1986; Di Benedetto, 1989; Huang, 1992; Sussman, 1990a). The goal is to establish (a) whether classification accuracy varies as a function of method and (b) whether all methods are equally appropriate for all vowel categories. The second part of the chapter considers alternative means of representing the frequency values and evaluates the classificatory performance of Syrdal & Gopal's (1986) bark-difference model of vowel-recognition against linear spectral distance measures and absolute frequencies.

4.1 Comparison of alternative sampling points

4.1.1 Review

The putative vowel target is usually equated with that part of the vowel where the contextual influence is minimal. This is generally assumed to be the central section, hence the widespread use of the durational vowel midpoint as a means of representation. However, evidence suggests that vowels vary inherently with regard to the slope and duration of transitions into and out of the vowel nucleus (Peterson & Lehiste, 1961; Stevens, House & Paul, 1966; Strange et al., 1983). Therefore, measures such as the steady-state and formant extrema values, which are not temporally fixed, may be more accurate methods of characterisation.

Lindblom (1963) defines the vowel target as the turning point in the trajectories of the first three formants. Other researchers sample the formants at the turning point for either F1 (Lisker, 1984; Di Benedetto, 1989; Veatch, 1991) or F2 (Huang, 1992). One argument for selecting the F1 extremum is that this is likely to be the most perceptually salient point in the vowel since it reflects the greatest degree of jaw opening and should therefore also coincide with the point of maximum amplitude (Huang, 1992; Veatch, 1991).

In a comparative study of different sampling points, Huang (1992) reports a slight improvement in classification accuracy for a combination of the F2 and F3 extrema with the mean value for F1 across the middle 50% of the vowel than for values sampled at either the F1 extremum or at the midpoint. This finding accords with evidence from perceptual studies that F1 is perceived with temporal averaging (Huang, 1985; Di Benedetto, 1987) while F2 is perceived with overshoot (Lindblom & Studdert-Kennedy, 1967). Lindblom & Studdert-Kennedy presented listeners with synthetic /jVj/ and /wVw/ syllables. They found that the F2 values at which the percept of the vowel changed from the back vowel /ʊ/ into the front vowel /ɪ/, was higher in the case of the /j-j/ frame than in the case of the /w-w/ frame. The listeners thus reported higher values than were actually reached in the stimuli, indicating perceptual overshoot.

These measures may be well motivated with respect to what is known about production and perceptual strategies but their usefulness as general descriptors of vowel quality is questionable on account of their limited practical applicability. For example, depending on the properties of the vowel in question and its adjacent context, the formant trajectory in a CVC sequence may not display any deviation from the linear transition between adjacent consonants that can be associated with specific movement towards the vowel target. This may be due, on the one hand, to extreme consonant-vowel coarticulation (Krull, 1989) or, on the other, to the fact that no deviation is required since the vowel target represents an intermediate point on the consonant-to-consonant trajectory. Alternatively, the trajectory may show considerable deviation from linearity but, depending on context, may show an "L-shape" rather than a "U-shape". These alternative possibilities are illustrated schematically in Figure 4.1 where (a) represents U-shaped trajectories, (b) represents extreme consonant-vowel coarticulation and/or linear interpolation through and (c) represents L-shaped trajectories.

U-shaped trajectories are likely to be more common in the case of F1 than F2 since the F1 target values for vowels are typically higher than F1 consonantal locus values, resulting in CVC trajectories characterised by a maximum value toward the center and lower values at the consonant-vowel boundaries (Stevens, House & Paul, 1966; Di Benedetto, 1989). Vowels typically involve a more open constriction than consonants and an increase in jaw opening generally implies an increase in F1 frequency value. The F1 extremum may thus provide a more reliable measure than the F2 extremum in terms of being more widely applicable.

A potential problem with selecting either the F1 or the F2 extremum, is the possibility that vowels vary with regard to which formant is the most perceptually salient. Van Son (1993), for example, claims that the size of the F1 excursion is relevant to vowel identity in the case of low and mid-low vowels which are characterised by high F1 values but is less important for high vowels which are characterised by low F1 values. For this reason, use of the F1 extremum as a general sampling point may be inappropriate. Similarly, while formant values sampled at the midpoint may provide a reasonable representation of the vowel target in the case of tense vowels, they may be less representative of target values in lax vowels. Tense vowels are believed to be characterised by transitions of approximately equal slope and duration, whereas lax vowels have been shown to display considerable asymmetry in their transitions (Lehiste & Peterson, 1961; Strange et al., 1983). Given that the target is reached earlier and held for a shorter period of time in the case of lax vowels, sampling the formants at the temporal vowel midpoint may yield values characteristic of the vowel offglide rather than the target.

The question of whether or not different methods show a bias towards different vowels is of central concern in the present study. Since a goal of this thesis is to compare vowels with respect to the relative variability they exhibit as a function of context, it is important that the means of representation selected should be equally as appropriate for all the vowel categories under study. This is to ensure that differences between vowels with respect to the variability they exhibit is attributable to differences in their relative degree of sensitivity to context, not to what amounts to an incorrect parameterisation of the target in the case of some (but not all) vowels. This question also has important implications for the interpretation and comparison of results across studies which employ different measurement procedures.

In the following analysis, methods will be evaluated in terms of their practical applicability and the comparative success with which they characterise individual vowels as well as in terms of overall classification accuracy.

4.1.2 Methodology

F1 and F2 were sampled at (1) the durational vowel midpoint (Midpoint), (2) the vowel steady-state (Steady-state) and, for those trajectories showing a clear turning point at (3) their respective points of maximum displacement (F1/F2max), (4) the point of maximum displacement for F1 (F1 max) and (5) the point of maximum displacement for F2 (F2 max). To allow for the possibility that formants do not reach their targets at the same time, F1 maximum and F2 maximum/minimum values were also combined with F2 and F1 values averaged across the middle 50% of the trajectory, (6) F1 max, F2 mean-mid and (7) F2 max, F1 mean-mid. Values were also averaged for both F1 and F2 across (8) the middle 50% of the vowel (Mean-mid) and (9) the entire vowel duration (Mean).

Steady-state values for the vowel were obtained automatically using an adaptation of Van Bergem's (1988) formula which locates the point in the vowel where the rate of change in the logarithm of F1, F2 and F3 is at a minimum. Displacement extrema for U-shaped trajectories were also located automatically by calculating the amount of deviation shown by the maximum and minimum value in the trajectory from a straight line drawn between vowel onset and offset. The point which represents the most deviation from the line is regarded as the target.

Success in characterising vowels is measured in terms of the separability of vowel distributions. This is quantified statistically by means of a Bayesian classifier using a quadratic discriminant function (James, 1985). Quadratic discriminant analysis performs a classification function that assigns each sample token to a category on the basis of which estimated *a posteriori* probability of class membership is the highest.

A feature vector of formant values for each token serves as input to the classifier. The mean feature vector and covariance matrix is computed from the feature vectors for all the tokens within a given class. This serves as the class model. The discriminance between class models for each token is calculated. The individual quadratic discriminance scores (converted to negative logarithm base 2) represent the probability of correct class membership for each token. From these the mean negative log probability scores for each vowel class are calculated. These give an indication of the classification accuracy obtained for each vowel class. The 'entropy' of the classification is also determined (Shannon, 1949; Pierce, 1962). This provides a measure of the overall classification accuracy (i.e. across all vowel classes). The lower the mean negative log probability scores and the lower the entropy scores, the better the performance of the classifier. In the present data, the maximum entropy score possible is obtained when the probability of membership is equal for all classes, that is, when the log probability is log $(1/12)$. In this case, the entropy is equal to $-\log_2 (1/12) = \log_2{}^{12} = 3.58$.

Other researchers using discriminant analysis techniques to quantify vowel separability present results in terms of percent correct classification scores (Syrdal & Gopal, 1986; Di Benedetto, 1989; Sussman, 1990a; Huang, 1992). These represent the number of tokens within each class that are closer to the correct class model than to any other class model. Mean negative log probability scores provide a more accurate measure of the classifier's performance because they reflect the distance of all the tokens within a given class from the class model.

In order to gain an appreciation of how far the amount and direction of spread in frequency value varies across different methods, coefficient of variance values for F1 and F2 were also calculated for each vowel for each method. These are shown in Tables 4.1:4. The coefficient of variance expresses the standard deviation as a percentage of the mean. This normalises for differences in mean value as a function of formant frequency range and thus permits direct comparison of variability along F1 and F2 and of the variability shown by different vowels.

4.1.3 Results

4.1.3.1 Comparative discriminability of vowel tokens with U-shaped trajectories

Across all vowels, only 11% of tokens showed a clear turning point in the trajectory for both F1 and F2. A further 22% of vowel tokens were characterised by a turning point for F2 (but not for F1) while a further 23% of tokens were characterised by a turning point for F1 (but not for F2). The data is accordingly divided into four subsets: (a) F1/F2-max, those tokens for which turning points are discernible for both F1 and F2, (b) F1-max and (c) F2-max, those tokens in which either the first or second formant display turning points and (d) No-disp (no displacement), those tokens in which neither formant trajectory shows a clear turning point. The entropy scores obtained for each method within each subset are presented in Table 4.5.

Table 4.5 Entropy scores for each method within each vowel subset.
The scores for F1/F2 sampled at the F2/F1 point of maximum displacement are given in parenthesis.

	Midpoint	*Steady-state*	*F1/F2 max*	*F1 max*	*F2 max*	*Mid-mean*	*Mean*
F1/F max	1.03	1.07	1.12	1.08 (1.24)	1.05 (1.18)	1.02	1.27
F1 max	1.26	1.41	*	1.33 (1.40)	*	1.29	1.39
F2max	1.47	1.47	*	*	1.48 (1.68)	1.47	1.61
No disp	1.36	1.51	*	*	*	1.40	1.55

Wilcoxon signed rank tests were performed pairwise for each method across sets using the mean negative log probability scores obtained for each vowel category. These are given in Table 4.6:9. Results at a significance level of $p < 0.05$ indicate greater overall discriminability for tokens that are characterised by a U-shaped trajectory for both formants than for tokens which show a U-shaped trajectory for one formant only or which show no turning point in either trajectory. The F1/F2-max set shows greater overall classification accuracy than each of the other sets for each method except Mean where the score is not significantly different from the score obtained for this method in the F1-max set.

The differences in score between the F1-max set and the No-disp set are significant only in the case of the Mean method, indicating a poorer comparative performance for this method in the No-disp set. Differences between the F1-max and F2-max set are significant in all cases except for Steady-state. The Steady-state score also represents the only significant difference in recognition rate between the F2-max and the No-disp set. F2-max shows a significant improvement in score for this method compared with the equivalent score in the No-disp set.

4.1.3.2 Relative performance of the alternative representations

The relative performance of the different methods of representation is assessed firstly in terms of the classification accuracy obtained for individual vowels and secondly in terms of overall classification accuracy. Wilcoxon signed rank tests were performed pairwise across methods for each vowel within each subset using the negative log probability scores for individual tokens. Differences in classification accuracy between the Midpoint, Mean, Mean-mid and Steady-state methods for each vowel were also tested for significance using vowel data pooled across subsets. For this set, all non-steady-state tokens were excluded from the analysis in order to permit direct comparison of the four methods. Finally, Wilcoxon tests were performed using the mean negative log probability scores for each vowel category to determine which method, if any, gives the best overall performance.

In general, lower probability scores, and therefore higher classification accuracy, are obtained for the F1max, F2 mean-mid and F2max, F1 mean-mid representations than for the representations where both F1 and F2 are sampled at either the point of maximum displacement for F1 or F2. In the F1/F2-max set, the F1max/F2 mean-mid shows a significant improvement in classification accuracy compared with F1max for /a, ɔ, i, ɒ, ʌ/. Similarly, the score for F2max/F1 mean-mid is significantly better than F2max for /ɑ, ɔ, i, ɛ, ɒ, ʌ, u, ɜ/. Higher classification accuracy generally correlates with lower variances. With the exception of /ɜ, ə, u/, all vowels show higher variances for F2 values sampled at the F1 extremum than for the corresponding F2 mean-mid values. Similarly, all vowels apart from /ɜ/, show higher variances for F1 sampled at the F2 extremum than for the corresponding F1 mean-mid values. Wilcoxon results also demonstrate significant differences between the two representations in the F2-max set for all vowels except /ɒ/ and /ʌ/, and in the F1max set, for /i, ɪ, ɛ, ɔ, ə/.

Examination of the times at which F1 and F2 reach their respective displacement extrema in the F1/F2-max set shows that the extrema for both formants occur within the same time frame on only 12% of occasions and within 20 ms of each other in 46% of cases. (However, F1 and F2 both reach their maximum within the middle 50% of the vowel in 68% of cases.) Since the F1max and F2max measures obtain consistently poorer recognition rates than the F1/F2max, F2/F1 mean-mid measures, they are excluded from the remainder of the analysis. Henceforth F1max and F2max will refer only to the F1max, F2 mean-mid and F2max, F1 mean-mid combinations respectively.

The F1/F2max representation generally performs worse than either F1max or F2max. It obtains significantly poorer scores than F2max in the case of /a, ɜ, u, ɪ, ʌ, ʊ, ə/ and significantly poorer scores than F1max in the case of /i/, /ɑ/ and /ɒ/. This result is also attributable to higher variability for F1 maximum and F2 maximum/minimum values than for F1 and F2 mean-mid values and subsequently to a greater combined variance for the F1/F2 max representation. For example, the coefficient of variance values for /ɒ/ are 11.00 for F1 maximum values, 18.44 for F2 minimum values and 8.51 for F2 mean-mid values resulting in a combined variance of 19.51 and 29.44 for the F1max and F1/F2max representations respectively. F2 maximum/minimum values have the highest variance of all F2 measures for all vowels except /i/, /ɪ/, /ɛ/ and /ə/. F1 maximum values have the highest variance of all F1 measures for /ɜ, ɔ, u, ɛ, ʌ, ɒ, ʊ/, the highest variance after Mean for /a/ and the highest variance after Steady-state for /ə/.

Significantly better scores are obtained with the F1max than the F2max method for /ɑ/ and /i/ while F2max performs significantly better than F1max for /a/, /ɜ/, /ʊ/ and /ə/. In the case of /i/ the improved performance of F1max over F2max may be attributed to higher variances for F1 mean-mid than F2 mean-mid values and a consequently higher combined variance for F2max than F1max. /ɑ/ is the only vowel which obtains a significantly better recognition rate for F1max than F2max and lower variances for F1 maximum than F2 maximum/minimum values. Similarly, in the case of /a/ and /ʊ/, the improved performance of F2max is due to higher variances for F2 mean-mid than F1 mean-mid values. For both these vowels F2 maximum/minimum values are more variable than the corresponding F1 maximum values.

The apparent preference for F1max over F2max shown by /ɑ/ is also supported to some extent by the results in the F1-max and F2-max sets. For /ɑ/ tokens in the F2-max set the F2max representation performs significantly worse than any of the other methods whereas in the F1-max set, the F1max representation score for /ɑ/ is equivalent to that obtained by Midpoint and Mean-mid. In the case of /ʌ/ which is spectrally similar to /ɑ/, F1max obtains the joint best score with Midpoint. For the other vowels, F1max performs significantly less well than either Midpoint or Mean-mid.

In the F2-max set the F2max representation performs significantly better than the other methods in the case of /i/ and /ə/ and also performs comparatively well for /ɛ/. It performs significantly less well for /a/ than either Midpoint or Mean-mid and, after Mean, obtains the worst score for /ɔ/.

It was not possible to obtain steady-state values for 21% of /ə/ tokens, 11% of /ɪ/ tokens, 9% of /ʊ/ tokens, 3% of /u/ tokens and a limited number of /a, ɛ, i, ɒ, ʌ/ tokens (less than 1%). Because it is necessary to have matched pairs for the Wilcoxon tests, the statistical significance of differences in score between Steady-state and the full range of methods can only be tested in a few cases within the individual subsets. However, direct comparison of Steady-state with Midpoint, Mean and Mean-mid is possible within the pooled data set which excludes all tokens for which steady-state values were unobtainable. Within this set, Steady-state performs significantly worse than the other methods in the case of /a/, /ɔ/ and /ɒ/. It is joint worst with Mean for /ʊ/ and worst after Mean for /i, ɛ, ʌ, ʊ, ə/. It performs best for /ɑ/ and better than Mean-mid for /ɛ/.

The comparative performance of Steady-state is positively correlated with vowel duration. In the case of the F1-max and the No-disp tokens, which display the shortest mean durations, Steady-State performs comparatively poorly. It compares more favourably with the other methods in the F2-max set and in the F1/F2-max set which show significantly higher mean durations. For example, in the No-disp set, Steady-state displays the highest combined variance of all the measures for all the vowels apart from /a/ and /ɜ/. In the F1-max set it also shows higher variances than the other measures for /a, ɛ, ɑ, ɒ, ʊ, ɜ/. In the F2-max set, Steady-state obtains the highest recognition rate for /ɔ/ and /i/ and for /ɜ/ jointly with F2max and Midpoint. It also shows the lowest F2 variances for all vowels within this set and the lowest F1 variances in the case of /ɪ, a, ɜ, ʌ/.

In the pooled data set, Midpoint and Mean-mid give a significantly better performance than Mean and Steady-state in all cases. For the tense vowels /ɑ/, /ɔ/, /u/ and the long, lax vowel /a/, the scores for Midpoint and Mean-mid are not significantly different. Midpoint performs better than Mean-mid in the case of /i, ɜ, ɪ, ɛ, ʊ/ and schwa while Mean-mid gives the best performance for /ɒ/. Mean performs worse than other methods in the case of /i, ɜ, ɪ, ɛ, ʌ, ʊ/ and /ə /. For the other vowels it gives the worst performance after Steady-state.

4.1.3.3 Overall classification accuracy

In order to assess overall classification accuracy, a Wilcoxon signed rank test was performed pairwise across methods within each set using the mean negative log probability scores for individual vowel categories. In the No-disp set, the results reveal significant differences in score between each of the methods except Mean and Steady-state. The lower entropy scores obtained for the Midpoint and Mean-mid methods in this set, therefore, indicate greater overall classification accuracy with these methods than with either the Mean or Steady-state measures. In the F1/F2-max set, the score for Mean is significantly different from each of the other scores confirming that this gives the worst performance. There is also a significant difference in score between the Mean-mid and the F1/F2-max representations, showing a better performance for the former method. In the F2-max set, overall classification accuracy is similar for all methods except Mean which gives the worst result. In the F1-max set Midpoint gives the best overall classification accuracy followed by Mean-mid.

4.1.3.4 Providing additional information

Table 4.10 gives the entropy scores obtained for each method when duration is also given as an input parameter to the classifier.

Table 4.10: Entropy scores for each method within each vowel subset when duration is also given as an input parameter.

	Midpoint	*Steady-state*	*F1/F2 max*	*F1 max*	*F2 max*	*Mid-mean*
F1/F2 max	0.73	0.89	0.79	0.76	0.77	0.73
F1max	1.04	1.21	*	1.08	*	1.05
F2max	1.26	1.25	*	*	1.28	1.26
No disp	1.07	1.4	*	*	*	1.20

Percent correct classification scores computed for the Midpoint and Mid-mean methods in the pooled data set show an 8% improvement in overall classification accuracy with the addition of duration. They also show that this is entirely due to an improvement in the classification accuracy for the lax (short) vowels. The greatest improvement in score occurs for /ʌ/ (32%), /ɘ/ (19%) and /ɪ/ (10%). Not unexpectedly, durational information serves principally to reduce the confusion between tense/lax pairs. For example, in the case of /ʌ/ the greatest confusion is with /ɑ/. Without durational information, 38% of /ʌ/ tokens are misclassified as /ɑ/. When duration is supplied, this figure is reduced to 9%. The improvement in separability between /i/ and /ɪ/ is less at 2%.

4.1.3.5 Comparative discriminability of individual vowels

In addition to variation in score as a function of method, there is also considerable variation in classification accuracy as a function of vowel identity. In all sets the tense vowels and the long, lax vowel /a/ are consistently better recognised than the other lax vowels. For the No-disp tokens the mean score across methods for the tense vowels as a set is .90 compared with 2.00 for the lax vowels including schwa, and 1.86 excluding schwa. (The corresponding mean percent correct classification scores are 85%, 34% and 47% respectively.) Comparison of these figures with the corresponding figures in the F1/F2-max set, .75 (88%) for the tense vowels and 1.43 (46%) for the lax vowels including schwa and 1.20 (57%) excluding schwa, shows that the improvement in overall classification accuracy for the F1/F2-max set is due chiefly to an increase in the recognition rate for the lax vowels. Comparison of the mean scores for individual vowels shows that the largest improvement occurs for /ʌ/, /ɒ/ and /ʊ/. The F1-max and the F2-max sets also show lower mean scores for the lax vowels compared with the No-disp set 1.77 (1.60 excluding /ə/) and (1.78 excluding /ə/) respectively. The mean score for the tense vowels in the F1-max set is the same as that in the No-disp set and higher (1.05) in the F2-max set. Across all sets, the vowels that are consistently most well recognised are /a/, /i/ and /ɔ/. The vowels which consistently achieve the worst recognition rates are /ɪ/ and /ə/.

4.1.4 Discussion

4.1.4.1 The relative performance of alternative representations

The formant extrema measures proved to be the most limited in terms of practical applicability. Only 11% of tokens displayed a turning point in both formant trajectories. A higher percentage of tokens displayed turning points for either F1 or F2. As predicted, more tokens were characterised by F1 extrema than by F2 extrema. However, the total number of these was still considerably less than might have been expected (e.g. see Stevens, House & Paul, 1966; Di Benedetto, 1989; Huang, 1992).

The relatively small proportion of vowel tokens characterised by U-shaped trajectories in the present data reflects the wide range of contexts and the lack of restriction of segmental sequences. Many of the previous studies which characterise vowels in terms of formant extrema use phonemically symmetrical CVC syllables (Lindblom, 1963; Stevens, House & Paul, 1966; Lisker, 1984) whereas the present data includes phonemically asymmetrical CVC syllables as well as other VC and CV combinations.

Huang (1992) and Di Benedetto (1989) use contexts which are asymmetrical with respect to the place of articulation of the adjacent consonants. However, Huang's data was controlled with respect to the voicing characteristic of the intial consonant and the place and manner characteristic of the final consonant. Di Benedetto's speech material was controlled with respect to the voicing characteristic of initial and final consonant. Furthermore, Huang used only stressed vowels. Di Benedetto used monosyllables in carrier phrases.

Stressed vowels and vowels in monosyllables tend to be longer than unstressed vowels (Fry, 1965) and vowels in polysyllabic words (Port, 1981) and vowels also tend to be longer preceding voiced consonants than preceding voiceless consonants (Peterson & Lehiste, 1960; House, 1961). According to the Undershoot Hypothesis (Lindblom, 1963) shorter durations result in increased coarticulation. Size of formant excursion has been shown to be inversely related to amount of consonant-vowel coarticulation (Lindblom, 1963; Krull, 1989; Sussman, 1990a). Tokens with displacement extrema show longer mean durations than the tokens without extrema (although the ratio of stressed to unstressed vowels is similar across sets). Thus, greater consonant-to-vowel coarticulation may also account, in part, for the relatively low number of tokens with displacement extrema in the present data.

4.1.4.2 Bias towards individual vowels

There is some evidence that low vowels are better characterised by F1 extrema than by F2 extrema. F1max performs better than either Midpoint or Mean-mid in the case of /ɑ/ and /ʌ/. For the other vowels, it performs worse than either of these methods. However, it also gives a worse performance for the low, front vowel /a/ in the F1/F2-max set than the other methods and a worse performance than either Midpoint or Mean-mid in the F1-max set. The central vowels /ɜ/ and /ə/ appear to be better characterised by F2 extrema than the other methods.

In the introduction to this chapter, it was suggested that Midpoint may be biased towards tense vowels since there is evidence that these are characterised by symmetrical transitions whereas lax vowels are characterised by asymmetrical transitions (Lehiste & Peterson, 1961; Strange et al., 1983). In the present study, no difference in the comparative performance of Midpoint was found as a function of the tense/lax distinction.

It is possible that, in the present data, inherent differences between tense/lax vowels with respect to the time at which the target is reached, are obscured by temporal variations induced by context. However, it may also be the case that the differences in the formant structure of tense and lax vowels reported in the literature simply reflects greater context sensitivity in the case of lax vowels. In the present data, tense vowels were consistently well recognised whereas lax vowels obtained consistently poor recognition rates. This suggests that tense vowels are less contextually variable than lax vowels. The apparent symmetry of the transitions for tense vowels may therefore simply be indicative of a longer, more stable target area.

4.1.4.3 Overall classification accuracy

Overall, the formant extrema measures perform poorly in comparison with the other methods. This may be attributed to the relatively high variability of F1 maximum and F2 maximum/minimum values compared with the other F1 and F2 measures. The variability, in turn, is due to the inherent context-sensitivity of this measure. The times at which F1 and F2 reach their point of maximum displacement vary across contexts. F1 and F2 are also affected differentially by context.

These results indicate that, in absolute terms, values sampled at the formant extrema are no more meaningful for vowels in context than values sampled at other points in the formant trajectory. However, this does not necessarily imply that formant extrema are less relevant perceptually than other aspects of the trajectory. The results, to some extent, are an artefact of the evaluation procedure. Although greater separability in the distribution of acoustic values may assist vowel recognition, the listener also has recourse to dynamic contextual information as well as other acoustic information. It is, therefore, possible that these measures are more representative of the target than other measures if supplemented with temporal and relational information. For example, Di Benedetto (1989) reports greater classification accuracy for vowels along the height dimension when the slope and duration of transitions and the F1 onset frequency values were provided in addition to the F1 maximum value.

In terms of statistical separability, the Midpoint and Mean-mid give the best performance. In the F1-max and No-Disp sets Midpoint obtains the lowest entropy score. This is due to the improved performance of Midpoint over Mean-mid for lax vowels and the higher ratio of lax to tense vowels. Averaging across the entire trajectory results in low classification accuracy because transitional information is incorporated into the representation. Assuming that reduction is duration dependent (Lindblom, 1963), for short vowels, averaging across the middle 50% of the vowel is also likely to result in less contrast between vowels for the same reason. With the exception of /a/, lax vowels are inherently shorter in duration than tense vowels (Peterson & Lehiste, 1961). The tokens in the F1-max and the No-disp sets also show shorter mean durations than the F2-max and F1/F2-max sets.

4.1.4.4 Summary

In the present data, there are significant differences between vowels with respect to overall discriminability as measured in terms of the statistical separability of formant values. A tighter clustering of values and higher recognition rates are obtained for vowel tokens which are characterised by turning points in both formant trajectories than tokens which display a turning point in one formant only or which show no turning points in either trajectory. Classification accuracy also varies significantly as a function of vowel identity. Tense vowels are generally better recognised than lax vowels. /i/, /a/ and /ɔ/ obtain the highest recognition rates across methods while /ɪ/ and /ə/ are the most poorly recognised.

Classification accuracy also varies significantly as a function of method. Overall, Midpoint and Mean-mid give the best, and Mean the worst, classification accuracy. For the tokens which display extrema, F1max and F2max perform better than F1/F2max but generally perform worse than Midpoint and Mean-mid. The performance of Steady-state is variable across vowels and vowel subsets and is, predictably, also correlated with vowel duration.

There is some variation in the ranking of individual methods across vowels. However, the pattern of results for individual vowels is not constant across vowel subsets which suggests that there may be a context x method as well as vowel x method interaction. In order to test the general validity of these results, therefore, it would be necessary to repeat the comparison with other data. For present purposes, however, the only evidence of vowel bias is the comparatively better performance of F1max in the case of /ɑ/ and of F2max in the case of /ɜ/ and /ə/ and a generally poorer performance of Steady-state in the case of lax vowels than in the case of tense vowels.

Steady-state is not considered suitable as a general sampling point in view of its variable performance across vowels and subsets. The formant maxima measures are discounted largely on account of their limited practical applicability but also on account of their relatively poor performance in terms of overall classification accuracy. Of the remaining methods, Mean is discounted on account of giving the worst overall classification accuracy. Since Midpoint gives the best overall classification, this method is used in subsequent analyses.

4.2 Comparison of different representations of the formant fequency values

Non-linear auditory transforms and spectral distance measures have been shown to successfully normalise differences in frequency range across different speaker groups (Syrdal & Gopal, 1986; Hillenbrand & Gayvert, 1993). It is, however, not known whether these methods offer any advantage over linear and absolute frequency values in terms of normalising within speaker variability as a function of phonetic context. In order to address this question, Syrdal & Gopal's (1986) bark-difference model is compared with linear spectral distance measures and absolute frequencies to determine which gives the best classification accuracy for the vowels in the present data. The bark-difference model is selected for special attention in the present study since Syrdal & Gopal themselves suggest that it may serve to normalise rate- and context-dependent variability. They make this suggestion in the light of results from earlier studies (Syrdal & Steele, 1985; Syrdal, 1985) which indicate that binary feature classification along bark-difference dimensions is robust across different phonetic contexts.

4.2.1 Syrdal & Gopal's (1986) Bark-difference model of vowel recognition

In Syrdal & Gopal's (1986) bark-difference model of vowel recognition, vowels are represented in terms of the auditory distance between their component frequencies and between F1 and F0. Fundamental and formant frequencies are converted to a critical band (Bark) scale. The critical band scale divides the frequency range into twenty four perceptually equivalent units. It increases with frequency linearly to 500 Hz and approximately logarithmically thereafter. Once the values have been bark-transformed, F0 is subtracted from F1 and F2 is subtracted from F3 and the simple Euclidean distance between them represents auditory distance.

A critical distance of 3 bark for the "spectral center of gravity effect" (Chistovich & Lublinskaya, 1979) is used to classify vowels into binary categories along each bark-difference dimension. The spectral center of gravity effect refers to the auditory averaging of two or more formants. Perceptual studies with synthetic vowels indicate that

> the ear effectively averages two formants which are relatively close together ... and receives from them an overall quality roughly equivalent to that which would be produced by a single intermediate formant. (Delattre et al., 1952, p. 209)

On the basis of the results of these and similar studies, Chistovich & Lublinskaya (1979) suggest a model of auditory processing which consists of two stages: (1) spectral peak picking and (2) integration of these peaks over a critical distance of 3 to 3.5 Bark. Integration of F1 and F2, when it occurs (i.e. when the formants are within 3 to 3.5 Bark of each other), produces the phonetic percept of backness. Fant (1983) reports that all Swedish back vowels may be characterised by a distance between F2 and F1 of within the 3-bark critical distance whereas front vowels have values in excess of this distance.

Syrdal & Gopal (1986) report that the F2-F1 measure did not work so well for American English vowels as it had for the Swedish vowels and suggest that the auditory distance between F3 and F2 is a more reliable means of distinguishing front from back vowels. Front vowels are therefore characterised by F3-F2 bark-difference values of below 3 bark while back vowels have values which exceed this critical 3 bark distance.

In order to be able to distinguish vowels along the dimension of vowel height, Syrdal & Gopal expand the critical distance metric to include the auditory distance between F1 and F0. They claim that vowels may be classified as high or low depending on whether the distance between F1 and F0 falls below or above 3 bark: high vowels which are characterised by lower F1 and higher F0 values than mid and low vowels have F1-F0 bark-difference values within 3 Bark while mid and low vowels have F1-F0 difference values which exceed this distance. Despite the articulatory and acoustic differences between F0 and formant frequencies, Syrdal & Gopal argue for inclusion of F0 in the model on the basis that intrinsic pitch has been shown to vary systematically across vowels. Perceptual experiments have also shown F0 to influence phonetic judgements (Fant, Carlson & Granstrom, 1974; Traunmuller, 1981).

Syrdal & Gopal tested their model on the Peterson & Barney (1952) database which comprises vowels in /hVd/ contexts produced in citation-form by a group of seventy six male, female and child speakers. They report a significant improvement in vowel recognition across different speakers for bark-difference measures over linear frequencies. Evidence from earlier studies (Syrdal & Steele, 1985; Syrdal, 1885) indicates that binary feature classification is robust across different segmental and prosodic contexts despite variability in the bark-difference values themselves. In the light of this, Syrdal and Gopal suggest that binary feature specification along the two bark-difference dimensions may also serve to normalise rate- and context-dependent variability:

> Because of the extreme acoustic variability of speech, robust context- and rate-independent normalisation may best be achieved through phonetic feature classification. In this view, the transformation from acoustic to bark-difference values at the specific bark-difference level reduces variability between speakers, and the further transformation from bark-difference values to phonetic features at the critical bark difference level normalises within-speaker variability related to contexts and rate. (Syrdal & Gopal, 1986, p.1094)

Additional support for this is provided by Sussman (1990a) who reports high classification accuracy (95.7%) into front/back categories along the F3-F2 bark-difference dimension for vowels in /bVt/, /dVt/ and /gVt/ contexts.

4.2.2 Methodology

The fundamental and the first three formant frequencies were sampled at the durational vowel midpoint. The values were bark-transformed using the formula from Syrdal & Gopal (1986) and F1-F0, F3-F2 and F2-F1 bark-difference measures were calculated. Quadratic discriminant analyses were performed for (a) individual vowels and (b) front/back and high/low categories. As the majority of /u/ and /ʊ/ tokens in the present data are fronted in quality, these vowels are excluded from the front/back analysis. The results are compared with mean negative log probability scores obtained for corresponding linear spectral distance measures and absolute frequency values.

4.2.3 Results

4.2.3.1 Individual vowel classification

Table 4.10 presents the percent correct classification scores for individual vowels using F1-F0, F2-F1 and F3-F2 bark-difference measures. The percent correct classification scores are given in place of the mean negative log probability scores in order to facilitate comparison of results with Syrdal & Gopal's results which are shown in the first column of the table. Classification scores obtained in the present study for the F1-F0 and F2-F1 bark-difference dimension are also given.

Overall classification accuracy for the present vowel data is poor compared with Syrdal & Gopal's results, 46% compared with 82%. However, the discrepancy in score is chiefly due to low recognition accuracy in the present data for the lax vowels, 30% including schwa and 48% excluding schwa, compared with 82% for the tense vowels.

4.2.3.2 Front/back classification

Examination of the individual bark-difference values shows good separation into front/back categories along both the F2-F1 and F3-F2 bark-difference dimensions using a critical distance measure of 3.5 Bark (Chistovich & Lublinskaya, 1979). The critical distance of 3 Bark employed by Syrdal & Gopal did not fit the present data.

The percentage of front vowels which meet the critical distance metric with F2-F1 bark-difference values in excess of 3.5 Bark is near 100%. Only fourteen tokens, one /ɪ/ token and thirteen /a/ tokens have values within this distance. Classification of back vowels along the F2-F1 bark-difference dimension is less successful, 79% of tokens are characterised by bark-differences within 3.5 Bark: 7% of /ɑ/ tokens, 18% of /ɔ/ tokens, 12% of /ɒ/ tokens and 41% of /ʌ/ tokens have bark-difference values exceeding this distance. Syrdal & Gopal also report poor classification of /ʌ/ along this dimension. They were also unable to classify /u/ and /ʊ/ using F2-F1 bark-difference values.

The F3-F2 dimension is more successful in characterising back vowels. 98% of back vowels have values in excess of 3.5 Bark. There is, however, a decrease in the number of front vowels which meet the critical distance metric along this dimension; 89% of front vowels have values within 3.5 Bark. The reduced performance for front vowels is largely due to the failure to characterise /a/ (58 /a/ tokens (48%) exceed the critical distance).

The differences between the F2-F1 and F3-F2 measures in characterising front versus back vowels are reflected in the entropy scores. These are given in Table 4.11. Lower entropy scores are obtained for front vowels using the F3-F2 bark-difference values and for back vowels using the F2-F1 values. In terms of overall classification accuracy there is little difference between the two dimensions.

4.2.3.3 High/low classification

The F1-F0 bark-difference dimension is less successful in characterising vowels than the F3-F2 dimension. 92% of high vowels meet the critical distance criterion along F1-F0 with values below 3.5 Bark. However, only 60% of mid and low vowel tokens have F1-F0 values in excess of 3.5 Bark. Table 4.12 presents the entropy scores for classification into high/low categories using F1-F0, F2-F1 and F3-F2 bark-difference and linear spectral distance measures. As this table demonstrates, overall classification is poor. This is due chiefly to poor discriminability between vowels along F0. Inclusion of F0 serves to lessen distinctiveness along the height dimension between vowels. Better classification scores are obtained when bark-transformed F1 values are used in place of the F1-F0 bark-difference values.

4.2.3.4 Comparison with linear spectral distance measures and absolute frequencies

Table 4.14 presents the classification scores across vowels obtained for linear and bark spectral distance measures and absolute frequencies.

Table 4.14: Entropy scores for linear and bark spectral distance measures and absolute frequencies.

	F1-F0, F2-F1	*F1-F0, F3-F2*	*F1, F2-F1*	*F1, F3-F2*	*F1, F2*	*F1, F2, F3*
Linear	1.52	1.72	1.39	1.6	1.39	1.32
Bark	1.57	1.63	1.39	1.49	1.38	1.32

The linear spectral distance measure for F1-F0, F2-F1 performs better than the Bark equivalent. Overall, better scores are obtained for measures which combine F1 with F2-F1 or F3-F2 than for F1-F0 and F2-F1 or F3-F2. Linear F1, F2-F1 gives the best score for the spectral distance measures followed by Bark F1, F2-F1. For front/back classification, linear distance measures also perform better than the Bark difference measures. The best separation between front and back, however, is obtained for linear absolute frequencies.

4.2.3.5 Discussion

Results for individual bark-difference measures shows that the model fails principally on its high/low classification along the F1-F0 dimension. Mid and low vowels are not uniformly distinguished by F1-F0 bark-difference values. Although high vowels are characterised by F1-F0 values within the critical 3.5 Bark distance, it is F1 which serves as the distinctive parameter. Given that Syrdal & Gopal used acoustic data derived from the Peterson & Barney (1952) database which comprises isolated /hVd/ words, it is not inconceivable that intrinsic pitch differences across vowel categories should have functioned distinctively. However, in the present connected speech data it appears that any such differences are obscured by the modulations in pitch that occur as part of the sentence intonation and due to contextual influences (Silverman, 1987).

Better separation is achieved into front/back categories along either the F3-F2 or F2-F1 dimension. However, overall classification accuracy is considerably lower than that reported by Syrdal & Gopal. 28% of vowel tokens fail to meet the critical distance criterion along F3-F2 in the present data, compared with .008% of tokens reported by Syrdal & Gopal. Furthermore, in the present data, bark-difference measures do not perform any better than linear spectral distance measures or absolute frequencies in separating front and back vowels. This also contrasts with Syrdal & Gopal's findings that bark-difference measures resulted in a significant improvement in classification accuracy over linear as well as other non-linear measures.

The discrepancy in results between the two studies indicates that while bark-difference measures are successful in normalising variability that arises as a function of physiological differences between speakers, they are less successful in normalising contextual variability within the speech of a single speaker. Bark-difference measures are intended to represent the spatial patterns of excitation in the peripheral auditory system (Chiba & Kajiyama, 1941; Potter & Steinberg, 1950). The high classification accuracy reported by Syrdal & Gopal supports the hypothesis that these remain stable across speakers regardless of differences in absolute formant frequency range (Syrdal, 1985). However, this appears to be true only for vowels produced in /h-d/ contexts. In section 4.1 it was shown that the temporal relationship between F1 and F2 varies as a function of context as well as the absolute target values. The present results which, in contrast to Syrdal & Gopal's study, are based on data for a single speaker, suggest that the extent of the variability is such that binary feature classification along the bark-difference dimensions is not robust across different phonetic contexts.

4.3 General Summary

Alternative one-point representations of the formant trajectory were evaluated in terms of (a) overall classification accuracy, (b) vowel bias and (c) practical applicability. Results show a significant variation in classification accuracy as a function of method. In absolute terms, Midpoint values provide better classification than formant extrema or steady-state values. The F1max, F2max, F1/F2max and Steady-state methods are also restricted in their practical applicability by contextual and durational requirements. There was little evidence of bias in favour of different vowels. Although the ranking of methods varied across vowels, this also varied across vowel subsets which suggests a context x method interaction.

The classificatory performance of Syrdal & Gopal's (1986) bark-difference model of vowel recognition was also compared with the performance of linear spectral distance measures and absolute frequencies. Although bark-difference dimensions provide good front/back classification, they offer no advantage over linear values. High/low vowels were not uniformly distinguished along F1-F0. In view of these results, vowels in the present study will be represented by linear F1 and F2 frequencies sampled at the durational vowel midpoint.

Figure 4.1: Stylised formant trajectory shapes: (a) U-shaped, (b) Linear, (c) L-shaped

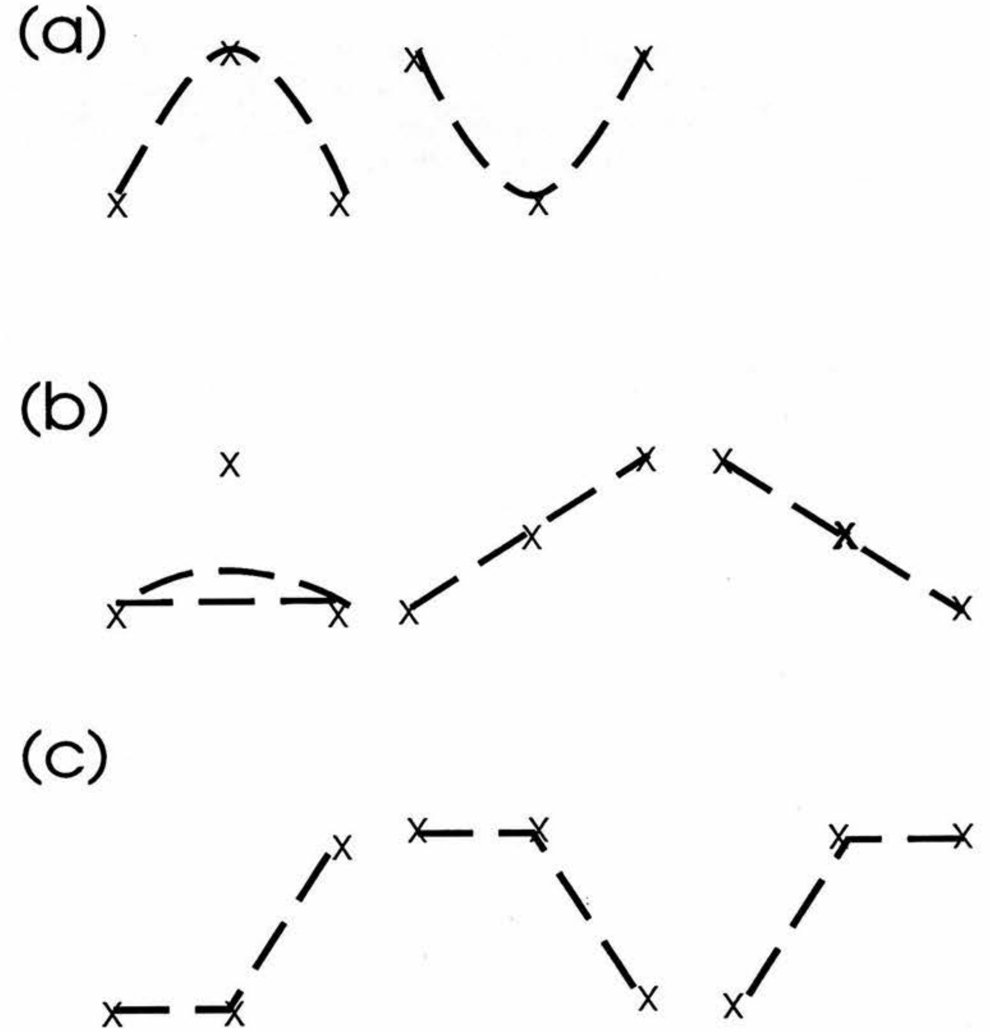


Table 4.1: Coefficient of variance values: No-disp set

	(1) Mid- point		*(2) S-S*		*(8) Mean- mid*		*(9) Mean*	
Vowel	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*
i	12.91	4.81	14.19	4.95	12.85	4.88	13.85	5.12
a	7.54	6.53	7.72	6.29	7.61	6.52	7.83	7.00
ɜ	4.63	5.86	5.37	6.03	4.86	6.20	6.02	7.23
ɑ	5.56	6.49	6.30	6.43	5.63	6.59	5.95	6.74
ɔ	7.69	12.29	9.36	17.63	7.44	11.83	7.52	11.59
u	9.31	16.73	8.83	19.79	8.98	16.43	8.77	15.53
ɪ	12.64	12.34	13.42	12.16	12.57	12.31	12.37	12.26
ɛ	9.35	9.92	10.85	10.24	9.23	10.01	9.52	10.32
ə	16.75	13.32	19.15	14.24	16.64	13.24	16.41	13.10
ʌ	9.97	10.71	11.39	12.40	9.82	10.70	9.32	11.05
ɒ	11.92	13.34	13.11	15.23	11.35	13.28	10.87	13.37
ʊ	8.90	19.82	8.89	21.75	8.73	19.62	9.18	19.10

Table 4.2: Coefficient of variance values: F1/F2-max set

	(1) Mid- point		*(2) S-S*		*(4) F1- max*		*(5) F2- max*		*(8) Mean- mid*		*(9) Mean*	
Vowel	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*
i	19.47	4.10	19.35	4.65	17.57	5.94	18.77	4.33	18.36	3.98	16.80	5.96
a	6.20	5.64	6.07	6.03	6.18	6.36	10.16	9.97	6.08	5.67	7.13	6.94
ɜ	7.14	4.91	7.77	4.77	10.74	4.50	6.82	8.66	6.74	4.74	6.84	5.75
ɑ	6.71	5.05	8.70	5.76	5.07	5.47	12.32	15.87	5.81	4.62	5.76	5.84
ɔ	6.57	6.22	7.05	5.98	9.42	16.25	7.99	21.80	5.65	6.52	5.65	9.95
u	12.61	6.35	12.66	6.41	15.96	5.93	13.29	8.47	13.06	6.26	12.51	6.40
ɪ	14.70	11.16	15.49	11.93	14.49	11.63	15.01	12.16	14.04	11.08	12.65	10.72
ɛ	8.04	8.73	8.36	8.57	9.13	9.28	11.53	9.97	7.65	8.86	8.23	10.01
ə	14.55	12.39	15.02	12.96	14.93	11.81	14.93	12.52	14.28	12.39	13.84	12.60
ʌ	7.50	5.73	7.84	5.96	7.99	7.72	8.67	14.23	7.20	6.17	7.57	8.46
ɒ	10.63	8.73	10.43	8.60	11.00	9.47	12.56	18.44	10.32	8.51	10.07	9.90
ʊ	3.97	19.20	4.75	18.28	4.89	17.44	4.63	20.00	3.49	18.77	2.39	19.59

Table 4.3: Coefficient of variance values: F1-max set

	(1) Midpoint		*(2) S-S*		*(4) F1 max*		*(8) Mean-mid*		*(9) Mean*	
Vowel	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*
i	12.47	4.72	11.56	4.54	15.08	5.90	12.08	4.81	11.68	5.25
a	5.52	5.59	5.76	6.38	5.49	5.64	5.43	5.59	5.97	6.20
ɜ	6.60	5.12	7.12	5.99	7.42	4.95	6.31	5.58	6.18	6.48
ɑ	6.73	7.55	9.47	8.37	6.82	7.78	6.45	7.54	6.92	7.10
ɔ	7.44	14.33	7.91	15.82	8.38	15.03	7.08	14.73	6.93	14.71
u	7.67	14.80	5.22	19.51	9.67	14.54	7.49	14.98	6.84	15.04
ɪ	13.13	13.47	13.99	13.85	13.98	13.66	13.01	13.57	12.34	13.55
ɛ	8.01	9.74	9.84	11.16	8.96	10.42	8.14	9.85	8.34	10.01
ə	14.46	12.63	15.77	13.38	14.95	12.55	14.47	12.78	13.99	13.07
ʌ	8.23	8.08	10.87	11.56	8.70	7.34	8.30	8.20	9.05	8.44
ɒ	8.57	8.75	9.01	11.91	8.55	9.45	8.18	8.74	8.41	9.94
ʊ	10.32	24.83	9.95	26.36	11.04	23.27	10.15	24.76	9.53	23.32

Table 4.4: Coefficient of variance values: F2-max set

	(1) Mid- point		*(2) S-S*		*(5) F2- max*		*(8) Mean- mid*		*(9) Mean*	
Vowel	*F1*		*F1*	*F2*	*F1*	*F2*	*F1*	*F2*	*F1*	*F2*
i	11.43	4.66	4.63	4.44	4.73	11.93	11.13	4.71	11.23	5.29
a	6.32	5.86	6.06	3.31	9.57	14.50	6.23	6.05	6.59	7.21
ɜ	9.40	6.72	7.12	4.01	6.72	10.50	9.38	7.00	9.87	7.53
ɑ	10.20	6.77	7.20	3.59	9.91	18.21	9.88	6.18	8.03	5.76
ɔ	8.15	14.32	14.24	4.89	14.98	14.06	7.87	14.02	8.43	13.62
u	9.00	15.02	14.59	3.72	15.37	10.39	9.13	14.87	9.30	14.76
ɪ	14.43	12.15	12.42	4.72	12.84	15.66	14.38	11.99	14.04	11.30
ɛ	10.15	10.89	11.02	4.34	12.16	12.97	9.53	10.80	9.56	11.32
ə	17.71	12.08	12.47	6.60	12.47	17.60	17.63	11.93	17.36	11.72
ʌ	13.16	11.17	10.69	11.85	12.32	14.89	11.46	10.36	11.10	9.59
ɒ	10.75	10.19	10.42	4.78	10.91	14.06	10.50	10.00	10.22	10.53
ʊ	8.78	21.99	22.42	6.33	22.81	9.60	8.55	21.45	8.17	19.26

Table 4.6: Mean negative log probability scores: No-disp set

Mean negative log probabilities for each method and the overall mean probability score (and SD) across methods are given for each vowel. Figures are rounded to 2 decimal places.

	i	a	ɜ	ɑ	ɔ	u	ɪ	ɛ	ɒ	ʌ	ʊ	ə
Midpoint	0.65	0.49	0.88	0.79	0.79	1.29	1.98	1.22	1.78	1.94	1.84	2.68
S-S	0.78	0.62	0.88	1.03	1.03	1.39	1.94	1.49	2.00	2.30	1.87	2.84
Mid-mean	0.59	0.50	1.00	0.89	0.77	1.29	2.01	1.27	1.81	2.04	1.81	2.71
Mean	0.78	0.57	1.39	1.11	0.78	1.32	2.10	1.53	1.94	2.28	1.99	2.80
MEAN	0.72	0.55	1.04	0.93	0.84	1.32	2.01	1.38	1.88	2.14	1.88	2.76
SD	0.07	0.06	0.13	0.13	0.12	0.05	0.07	0.15	0.10	0.18	0.08	0.08

Table 4.7: Mean negative log probability scores: F1/F2-max set

	i	a	ɜ	ɑ	ɔ	u	ɪ	ɛ	ɒ	ʌ	ʊ	ə
Midpoint	0.52	0.54	0.90	0.98	0.49	0.96	2.03	0.99	1.07	0.93	0.39	2.51
S-S	0.48	0.58	0.92	0.99	0.42	0.92	2.14	1.04	1.28	1.01	0.50	2.55
F1/F2 max	0.52	0.60	1.19	0.80	0.43	1.12	2.15	1.12	1.10	1.33	0.47	2.57
F1max	0.47	0.60	1.11	0.77	0.46	1.09	2.07	1.13	1.04	1.06	0.59	2.56
F2max	0.55	0.54	0.76	1.01	0.38	1.00	2.13	1.01	1.13	1.34	0.30	2.48
Mid-mean	0.50	0.54	0.84	0.91	0.42	0.98	2.05	1.02	1.05	1.10	0.36	2.49
Mean	0.61	0.69	1.13	1.07	0.66	1.09	2.20	1.37	1.20	1.97	0.60	2.61
MEAN	0.52	0.58	0.98	0.93	0.47	1.02	2.11	1.10	1.12	1.25	0.46	2.54
SD	0.05	0.06	0.16	0.11	0.09	0.08	0.06	0.13	0.09	0.35	0.11	0.05

Table 4.8: Mean negative log probability scores: F1-max set

Mean negative log probabilities for each method and the overall mean probability score (and SD) across methods are given for each vowel. Figures are rounded to 2 decimal places.

	i	a	ɜ	ɑ	ɔ	u	ɪ	ɛ	ɒ	ʌ	ʊ	ə
Midpoint	0.64	0.46	0.92	1.26	0.89	0.99	1.97	1.00	1.30	1.68	1.50	2.50
S-S	0.61	0.62	1.05	1.28	1.08	0.80	1.95	1.31	1.69	2.11	1.79	2.58
F1max	0.70	0.47	1.05	1.21	0.88	1.26	2.15	1.09	1.26	1.67	1.66	2.59
Mid-mean	0.68	0.46	0.96	1.25	0.87	1.00	2.02	1.05	1.29	1.74	1.57	2.56
Mean	0.77	0.49	1.11	1.32	0.81	1.07	2.15	1.25	1.32	1.98	1.72	2.69
MEAN	0.68	0.50	1.02	1.26	0.91	1.02	2.05	1.14	1.37	1.84	1.65	2.58
SD	0.06	0.07	0.08	0.04	0.10	0.16	0.10	0.14	0.18	0.20	0.12	0.07

Table 4.9: Mean negative log probability scores: F2-max set

	i	a	ɜ	ɑ	ɔ	u	ɪ	ɛ	ɒ	ʌ	ʊ	ə
Midpoint	0.59	0.49	1.55	1.29	1.08	1.14	2.00	1.44	1.42	2.27	1.72	2.68
S-S	0.58	0.65	1.45	1.20	1.03	1.18	1.98	1.53	1.31	2.09	1.84	2.77
F2max	0.58	0.55	1.54	1.44	1.08	1.14	2.00	1.40	1.49	2.26	1.67	2.66
Mid-mean	0.59	0.45	1.62	1.28	1.07	1.18	2.01	1.46	1.36	2.17	1.70	2.70
Mean	0.71	0.65	1.88	1.23	1.10	1.44	2.13	1.74	1.43	2.30	1.88	2.87
MEAN	0.61	0.56	1.61	1.29	1.07	1.22	2.03	1.51	1.40	2.22	1.76	2.74
SD	0.05	0.09	0.17	0.09	0.02	0.13	0.06	0.14	0.07	0.09	0.09	0.08

Table 4.11: Percent correct classification scores for bark-difference measures. Syrdal & Gopal's (1986) results are given in the first column.

	S&G	*F1-F0, F3-F2*	*F1-F0, F2-F1*
i	95%	88%	88%
a	87%	90%	87%
ɜ	94%	81%	82%
ɑ	89%	73%	82%
ɔ	80%	89%	89%
u	77%	47%	70%
ɪ	84%	39%	48%
ɛ	87%	64%	64%
ə	*	5%	10%
ʌ	89%	43%	37%
ɒ	*	78%	54%
ʊ	77%	51%	51%
All	86%	46%	50%
Tense	87%	82%	85%
Lax	84%	48%	51%
Lax excl. /ə/	*	30%	34%

Table 4.11: Entropy scores for front/back vowel classification

	F1-F0 & F2-F1			*F1-F0 & F3-F2*			*F1 & F2-F1*			*F1 & F3-F2*		
Scale	*Front*	*Back*	*All*	*Front*	*Back*	*All*	*Front*	*Back*	*All*	*Front*	*Back*	*All*
Linear	.15	.33	.24	.23	.32	.27	.15	.32	.24	.24	.32	.29
Bark	.18	.31	.25	.21	.28	.25	.19	.29	.24	.22	.30	.26

Table 4.12: Entropy scores for high/low vowel classification

	F1-F0 & F2-F1			*F1-F0 & F3-F2*			*F1 & F2-F1*			*F1 & F3-F2*		
	High	*Low*	*All*	*High*	*Low*	*All*	*High*	*Low*	*All*	*High*	*Low*	*All*
Linear	.49	.50	.49	.51	.53	.52	.48	.51	.50	.49	.53	.51
Bark	.52	.54	.53	.57	.56	.56	.52	.56	.54	.54	.57	.55

Chapter 5

Role of Context in Conditioning Vowel Quality

Introduction

This chapter examines the nature and relative magnitude of coarticulatory effects on schwa compared with the full vowels with reference to the question of whether or not schwa can be associated with an independent phonetic target. If schwa is targetless, its midpoint value should be entirely predictable from context[2]. It should make no independent contribution to the trajectory between adjacent context segments but is expected to be interpolated through. It should, therefore, also allow coarticulatory effects between adjacent segments and show comparable amounts of anticipatory and carryover coarticulation (see section 2.3.2.2 for a full discussion).

The following analysis is divided into three sections. In the first of these, the linearity of schwa mean first and second formant trajectories is evaluated in comparison with mean formant trajectories for the full vowels. Mean formant trajectories, obtained by plotting the mean onset, midpoint and offset values for vowels as a function of context, provide a useful summary of the data and give an immediate visual impression of the extent to which schwa can be said to be interpolated through. Vowels are also compared in terms of the overall range and the comparative range in formant frequency value they exhibit at vowel onset, midpoint and offset. In the second part, the overall degree of context-dependency shown by schwa and the full vowels and the relative magnitude of different

---

[2] It is noted that 100% prediction accuracy is unlikely owing to the fact that some random variability is inevitable.

coarticulatory effects is assessed in a series of multiple regression analyses. In the final section, a series of one-way and two-way analyses of variance with post hoc Tukey multiple comparison of means tests are performed in order to determine the degree of differentiation in vowel midpoint values as a function of individual contexts.

5.1 'Interpolation through': mean formant trajectories and range in formant frequency values

Stylised mean formant trajectories were obtained for each vowel in each context by plotting the vowel mean onset, midpoint and offset values as a function of the place of articulation of the preceding and following consonants in the case of F2 and as a function of both adjacent consonant place and manner of articulation in the case of F1. Measures of the deviation from linearity for each mean trajectory were calculated by fitting a regression line to the three data points (i.e. mean onset, midpoint and offset value) and summing the squared residuals. The total residual distance, as a measure of the goodness of fit of the data points to their associated line regression, provides a comparative measure of how far each mean trajectory deviates from the linear interpolation between consonants. For easier interpretability, the square root of the summed squared residuals is used as the final comparative figure. Coefficient of variance values for F1 and F2 onset, midpoint and offset values were also calculated in order to assess overall variability at each time point. The results for F1 and F2 are considered separately, beginning with F2.

5.1.1 F2 mean trajectories

In order to determine the systematicity of context effects and the extent of the interpolation through schwa, the /ə/ mean second formant trajectories were grouped according to the place of articulation of either the preceding (C^1) or following (C^2) consonant. Thus, Figure 5.1a shows the mean second formant trajectories for $C^1əC^2$ sequences where C^1=palatal irrespective of C^2 place of articulation. Conversely, in

Figure 5.1b, C^2=palatal irrespective of C^1 place of articulation. Similarly, Figure 5.2a:b shows /ə/ mean trajectories in the context of a preceding or following velar.

In the key for each graph, place categories are abbreviated as follows: palatal and palato-alveolar (Pal), velar (Vel), alveolar (Alv), dental (Den), labio-dental (Labd), labial (Lab), apical (Api) and labio-velar (Labv). The place categories include all manner classes (e.g. 'Lab' includes both the bilabial oral and nasal stops) and both voiced and voiceless consonants. With the exception of /l/ and /r/, which are classified as apical, consonants are distinguished solely in terms of their place of articulation. The consonants /l/ and /r/ are distinguished from the other alveolars because they involve special articulations (i.e. lateral and retroflex respectively). The affricates /ʧ/ and /ʤ/ are included in the palatal category when they occur preceding vowels (since it is the palatal part that is immediately adjacent to the vowel) and in the alveolar category when they occur following vowels (since the alveolar part is immediately adjacent to the vowel) although, in effect, both parts are likely to show a palato-alveolar place of articulation. C^1 and C^2 contexts represent the immediately adjacent consonants either preceding or following the vowel irrespective of syllable- or word-boundaries. Thus the palatal and labio-velar glides /j/ and /w/ are included in the C^2 contexts despite the fact that, in English, these consonants only occur in pre-vocalic position. Similarly, the velar nasal /ŋ/ which only occurs post-vocalically, is included in the C^1 contexts. In each of these cases, /j, w/ or /ŋ/ represent C^2 or C^1 consonants in $V{*}C^2V$ or $VC^1{*}V$ sequences where * denotes a syllable- or word-boundary.

Comparison of the two groupings within each figure reveals a tighter clustering of offset compared with onset values (see section 7.6.2 for discussion). However, both groupings convey similar information with respect to the linearity of the trajectories. Thus, for the remaining contexts, only the schwa mean trajectories plotted as a function of C^1 place of articulation are shown (see Figure 5.3). (The corresponding trajectories plotted as a function of C^2 place of articulation are provided for reference in Appendix A.). The mean values and deviation measures are shown in Table 5.1. The corresponding values for the full vowels are given in Table A.1.a:k, Appendix A. Where there are no examples of a given vowel-context combination, this is indicated by a series of asterisks.

In the case of the full vowels, the mean second formant trajectories were divided into two broad categories according to whether they show a high or a low degree of deviation from linearity. For this purpose, a benchmark of 75 Hz was selected. Given the assumption that shifts in F2 below the order of 75 Hz are not perceptible (Flanagan, 1955), it is reasonable to expect deviation values of upwards of 75 Hz for targeted vowels, at least in those contexts which require a warping of the trajectory to produce a vowel percept[3]. In contexts where the vowel target lies intermediate on the consonant-to-consonant trajectory, low deviation values are expected. For example, interpolation through /i/ is expected in the context of segments which share the specification [+front] and are consequently also characterised by high F2 values (e.g. palatals and alveolars). Conversely, interpolation through /ɔ/ is expected between segments also characterised by low F2 values (e.g. labials, velarised /ɫ/ and labio-velar /w/).

Degree of deviation from linearity is expected to increase with an increase in the 'acoustic distance' between the vowel target and consonant locus value (Stevens & House, 1963). Thus, formant trajectories for /i/ should show most deviation from linearity in the context of a preceding or following /w/ while the trajectories for /ɔ/ are likely to show most deviation in the context of an adjacent palatal or alveolar. This pattern is evident in the present data and is illustrated in Figures 5.4 and 5.5 which show mean formant trajectories for /i/ and /ɔ/ respectively. (See Table A 1.a and 1.i, Appendix A, for the corresponding mean values and deviation measures.)

A higher percentage of low deviation values is predicted for schwa than for the full vowels on account of the fact that there are a greater number of contexts which involve the tongue trajectory passing through the central area than through the more peripheral

[3] This represents a generous estimate given that significantly larger difference limens have been reported for vowels in context (Mermelstein, 1978).

regions in the vowel space. However, low deviation values are also obtained for schwa mean trajectories in contexts where a significant warping of the trajectory would be expected if schwa were targeted. These are the contexts which yield low deviation values for the peripheral vowels and which are symmetrical with respect to the high/low F2 locus classification (i.e. where both the preceding and following consonants are characterised by either a high or by a low F2 frequency value). As Figure 5.6 demonstrates, in the present data, there is no warping towards a more central value in these contexts, rather, in each case, there is a straight-line interpolation from onset to offset through the schwa midpoint.

There are only four contexts (7% of the total) in which the schwa mean trajectory is characterised by a deviation value > 75 Hz. These are shown in Figure 5.7. In three of these: Lab-Pal, Vel-Lab and Pal-Labv, the mean midpoint value approaches the overall mean (and ostensible target) value for schwa of 1447 Hz. In the remaining context: Api-Labv, the mean midpoint value is considerably lower than the overall mean.

For comparison, the mean second formant trajectories for /ɜ/ and /a/ are presented in Figures 5.8:5.9. Both vowels display a relatively high percentage of low deviation values compared with the other full vowels on account of their more central articulation. As Figure 5.8b shows, /ɜ/ patterns with schwa insofar as lower deviation values are generally obtained in contexts which are asymmetrical with respect to the high/low locus classification although it is also interpolated through in the Alv-Labd, Den-Alv and Labd-Den contexts which are characterised by similar F2 values to its target value (defined here as its overall mean of 1379 Hz). However, in contrast to schwa, a clear warping of the trajectory is evident in symmetrical high and low locus contexts (see Figure 5.8a). Similarly, in the case of the low, front vowel /a/ it is possible to identify a relatively stable target area which the majority of mean trajectories either pass through or move towards (see Figure 5.11a). With the exception of one example of extreme overshoot in the Vel-Vel context (1671 Hz compared with an overall mean of 1492 Hz) and three examples of undershoot in the context of a following dark /ɫ/ (mean 1346 Hz), /a/ also displays a relatively tight clustering of mean midpoint values.

The mid-high, lax vowels /ɪ/ and /ʊ/ both display a comparable range in F2 mean midpoint value to schwa. This is illustrated in Figure 5.10 which shows the total number of mean trajectories for each vowel. In contrast to schwa, a relatively high percentage of the F2 mean formant trajectories for /ɪ/ show high deviation from linearity. However, there are only six examples which show a clear warping at the midpoint which may be construed as movement toward a target value. These are shown in Figure 5.11a. For the majority of the remaining cases, the mean midpoint value is similar (within 20 Hz) to the mean onset or mean offset value. The high deviation value, in these instances, results from either a steep offglide or onglide. For example, in the Vel-Vel, Pal-Vel and Vel-Alv contexts, F2 starts high, remains high at the midpoint and then falls sharply to the value at offset. In the Lab-Labd context, F2 starts low, remains low at the midpoint and then rises steeply to the value at offset. The mean trajectories in Figure 5.11c show a similar pattern except that, in these cases, the difference in mean onset/offset and mean midpoint value is greater (between 45 Hz and 87 Hz except for Pal-Api at 166 Hz). The same situation holds for /ʊ/ (see Figure 5.12). There is no narrowing in the range of mean midpoint value for either vowel across the high deviation contexts.

Coefficient of variance values calculated for each vowel across all tokens and shown in Table 5.2 and Figure 5.13, confirm higher levels of variability in F2 midpoint value for schwa, /ʊ/ and /ɪ/ than for all the other vowels with the exception of /ɯ/. More importantly, as Figure 5.13a demonstrates, also excepting /ɯ/, schwa, /ʊ/ and /ɪ/ show considerably less difference in amount of variability at the midpoint compared with at the onset and offset than the other vowels.

5.1.2 F1 mean trajectories

The F1 mean onset, midpoint and offset values and measures of deviation from linearity for schwa are presented in Table 5.3. Mean values and deviation measures for the full vowels are given in Table A 2.a:k, Appendix A. In the graphs pertaining to this section, place categories are abbreviated as before and manner categories as follows: 'st' (stop), 'n' (nasal), 'fr' (fricative) and 'gli' (glide). Manner is only specified where, for a given place of articulation, there is a manner contrast. Thus, the liquids and the labio-velar glide /w/ continue to be represented solely by their place category labels of 'Api' and 'Labv' respectively.

In the literature, lower overall variability and smaller and less systematic coarticulatory effects are reported for F1 than for F2 (Stevens & House, 1963; Koopmans van Beinum, 1994; Van Bergem, 1994). Less variation in F1 trajectory shape is also noted (Stevens, House & Paul, 1966, Di Benedetto, 1989). Stevens, House & Paul claim that since the F1 target values for vowels are typically higher than consonantal F1 locus values, the F1 trajectory in a CVC syllable is likely to be characterised by a maximum value towards the center and lower values at the CV boundaries irrespective of consonantal context. Contrary to this prediction, the first formant trajectories for vowels in the present data do not have the same form across all contexts or across all instances of a given context, owing to considerable variation in onset and offset frequencies (see also section 4.1.3.1). However, despite this variability, expected differences between vowels with respect to overall F1 excursion size do pertain. For example, as Figure 5.14 demonstrates, the majority of mean F1 trajectories for the low, front vowel /a/ and the low back vowels /ɑ/ and /ʌ/ show high deviation from linearity, where this is defined as being > 50 Hz, the difference limen for F1 as reported by Flanagan (1955). In contrast, the majority of mean trajectories for the high, front vowel /i/ are characterised by deviation values < 50 Hz. These are shown in Figure 5.15a.

Although characterised by lower absolute F1 midpoint values, the mid-low vowels /ɛ/, /ɒ/ and /ɔ/ pattern with the low vowels, displaying a higher proportion of high compared with low deviation values (see Figure 5.16). In contrast, schwa, also classified phonologically as a mid vowel, displays a higher proportion of low compared with high deviation values as illustrated in Figure 5.17. The high deviation mean trajectories are also plotted separately according to trajectory shape in Figure 5.18. As shown here, in the majority of cases, deviation values > 50 Hz are accounted for by either a steep onglide or offglide to or from a higher or lower value at onset or offset. There are only nine contexts where the mean F1 trajectory is characterised by a maximum value at the midpoint which might be interpreted as movement toward a vowel target. However, in each case, the preceding consonant is either /m/ or /n/. The large excursion at the midpoint of the trajectory in these cases, therefore, most likely reflects measurement errors introduced by the presence of nasal resonances and anti-resonances in the vowel spectrum rather than increased jaw opening in this context. There are three contexts, where the trajectory is characterised by a minimum value at the midpoint. The mean midpoint value of 449 Hz for these tokens is slightly lower than the overall mean.

The relatively high proportion of linear mean F1 trajectories and the wide range in F1 mean midpoint value evident for schwa, is further exemplified in a comparison of the schwa data with the data for the long, central vowel /ɜ/ (see Figure 5.20). In contrast to schwa, a relatively high number of the /ɜ/ mean trajectories show clear warping at the midpoint indicating movement towards a vowel target. With the exception of two examples of undershoot in the Alv.fr-Alv.fr and in the Alv.n-Labd contexts (449 Hz and 491 Hz respectively compared with an overall mean of 552 Hz) and one example of overshoot in the Lab.st-Api context (620 Hz), /ɜ/ also displays a comparatively narrow range in mean midpoint value across both high and low deviation trajectories (see also Figures 5.21 and 5:23).

In section 5.1.1 the mean F2 trajectories for the mid-high, lax vowels /ɪ/ and /ʊ/ were shown to display little deviation from linearity in the majority of contexts and to cover a comparable range in F2 to the range in F2 for schwa, suggesting that, along this formant dimension, these vowels are also interpolated through. In order to establish whether the same observation applies in the case of F1, the total number of mean F1 trajectories for /ɪ/ and /ʊ/ are plotted alongside the total number of schwa trajectories for comparison in Figure 5.25. The mean trajectories for /ɪ/ and /ʊ/ are also shown separately in Figures 5.26:5.27. In each of these figures, the trajectories are grouped according to whether they are characterised by high or low deviation values (i.e. greater or less than 50 Hz).

Comparatively little deviation from linearity is expected for /ɪ/ and /ʊ/ by virtue of their 'mid-high' classification. However, given the wide variability in onset and offset value, certain conditions should give rise to instances of displacement at the midpoint. For example, assuming that /ɪ/ is characterised by a target in the region of 360 Hz (Wells, 1962), a warping in the trajectory towards this value should be evident on those occasions where F1 has a higher frequency value at the vowel onset and offset than the putative target frequency. While there are clear examples of this in the case of /i/ (see Figure 5.15b), there are only two instances where this occurs for /ɪ/. The reverse situation, where the trajectory is characterised by a maximum value at the midpoint, applies in twelve contexts. As in the case of schwa, these are predominantly nasal contexts involving a preceding /m/ or /n/ and, therefore, most likely reflect measurement errors. With the exception of three examples where the mean trajectory is characterised by either a steep on- or off-glide, the remaining mean trajectories (as expected) show little deviation from linearity (see Figure 5.28.).

Overall, /ɪ/ displays a similar range in F1 mean midpoint value to that displayed by /ə/ (295 Hz compared with 310 Hz). However, the range for /ʊ/ is much narrower by comparison (153 Hz). As in the case of F2, the coefficient of variance values for F1 (see Table 5.2) confirm the high variability displayed by /ə/ and /ɪ/ compared with the other vowels and the comparatively little difference in overall range of value between onset, midpoint and offset values (see also Figure 5.13b).

5.1.3 Summary

In terms of the linearity of mean formant trajectories and the comparable spread in mean onset, midpoint and offset value, the data support the theory that schwa is phonetically transparent along F2 and F1. The present data also indicates a high degree of underspecification along both formant dimensions for /ɪ/ and along F2 for /ʊ/. In the following sections it will be shown that these vowels also display a relatively high degree of overall context dependency compared with the other full vowels and, in the case of /ɪ/, a similar patterning to schwa with respect to the systematicity and directionality of effects.

5.2 Multiple Regression Analysis

5.2.1 Introduction

Multiple regression/correlation analysis is a highly versatile data-analysis system that is particularly suitable for observational (non-experimental) research in which the parameter of interest is likely to be influenced by a multiplicity of inter-correlated factors (see Cohen & Cohen, 1983). In addition to providing a measure of the total variance in the dependent variable that is accounted for jointly by a set of predictor variables, it also permits a partitioning of the variance according to the unique contribution made by each predictor variable, relative to what is accounted for by the other predictors.

Multiple regression analysis is used here to evaluate the overall degree of context dependency shown by different vowels and the relative strength of carryover compared with anticipatory and of vocalic compared with consonantal coarticulatory effects. Vowel duration and the potential interaction of duration and context effects are also considered. It is assumed that degree of context-dependency is directly proportional to the amount of variance in midpoint values that is explained jointly by the predictor variables. The statistical package used was BMDP 1R. This fits the regression model:

$y = \alpha + \beta_1x_1 + \beta_2x_2 + \dots \beta_px_p + \varepsilon$, where y is the dependent variable, $x_1, \dots, x_p$ are the independent variables, $\beta_1, \dots, \beta_p$ are the regression coefficients, α is the intercept, p is the number of independent variables, and ε is the error. The null hypothesis predicts that when all the independent variables save one are held constant, the variable under investigation accounts for no y variance in the population.

Two series of analyses were performed. In the first, tokens within each vowel category are sub-categorised according to the particular sequence of consonants and vowels in which they occur, i.e. whether they occur in a VC-CV, CV-CV or VC-VC sequence. Each subset is considered separately. The independent or predictor variables include (1) preceding consonant (C^1), (2) preceding vowel (V^1), (3) following consonant (C^2), (4) following vowel (V^2), (5) schwa duration, (6) duration of V^1, and (7) duration of V^2. In the second series, tokens within each vowel category are pooled across all contexts. The predictor variables are simply (1) preceding context, (2) following context and (3) duration. For both series, the dependent variable is F1 or F2 sampled at the temporal midpoint of the target vowel. Consonants are characterised by frequency values averaged across the initial or final 10 ms of the adjacent vowel. Context vowels are characterised by their midpoint values. For the CV-CC and CV-CV sets, an additional analysis was performed using V^1 offset values to characterise V^1. This was to permit comparison of relative consonant versus vocalic influence in this temporal position.

5.2.2 Separate set analyses

The results for schwa are considered first. These are presented in Tables 5.4:5. Multiple R square (R^2) represents the proportion of the total variance in the dependent variable that is accounted for jointly by the predictor variables. The standardised partial regression coefficients (β values) indicate the relative weighting of the unique contributions of individual predictors. Following standard convention, squared correlation coefficients are rounded to four places and the standardised regression coefficients to two places (Cohen & Cohen, 1983). Because, in every case, the overall

R^2 value was significant at $p < .01$, significance levels are only shown for the individual β values. A double asterisk (**) denotes significance at $p < .01$ and a single asterisk (*), significance at $p < .05$.

5.2.2.1 Total variance explained

The proportion of total variance in F2 explained jointly by the full set of predictor variables ranges between 69% and 94%. The highest percentage of explained variance is obtained in the CCəCV subset. The two worst performances occur for tokens in the CVəCC and CVəCV sets where V^1 is characterised by values sampled at the V^1 midpoint. The performance of V^1 as a predictor variable in these sets improves considerably if V^1 is characterised by its frequency value at vowel offset.

The relatively poor performance of V^1 midpoint values (V^1mid) compared with the V^1 offset values (V^1off) is attributable to the high correlation between C^1 and V^1mid. C^1 is not correlated with schwa but it is highly correlated with V^1mid (.854 for the CVəCV set and .806 for the CVəCC set). This adds irrelevant variance to V^1mid and thereby weakens its relationship with schwa. When C^1 is partialled from V^1mid, the unique variance of V^1mid in schwa is greater than the proportion it accounts for when it is considered alone. For example, in the CVəCV set, the simple univariate correlation between V^1mid and schwa is .517 with an r^2 of .2673 ($p < .01$). When C^1 and V^1mid are analysed together there is a marginal increase in the total amount of variance explained, $R^2 = .2820$, but a substantial increase in the standardised regression coefficient for V^1mid, .72 ($p < .01$). This compares with -.23 for C^1 which fails to meet the .05 significance criterion. C^1 also acts to suppress V^1off but to a much lesser extent. The univariate correlation for V^1off in the same set is .872, r^2=.7617. With C^1 also considered in the equation R^2=.7738 with β values of .13 (p =.09) and .94 ($p < .01$) for C^1 and V^1off respectively. C^1 and V^1mid are also highly correlated for F1, r=.766 and r=.769 for the CVəCC and CVəCV sets respectively. Again, an increase in the standardised regression coefficient for V^1mid when C^1 is partialled out indicates that the effect of V^1 is being suppressed by C^1. Henceforth, results will only be considered for V^1off referred to as V^1.

The proportion of explained F1 variance is generally less than the proportion of explained F2 variance, ranging between 63%-88% (mean 75%). The range for F2, excluding the CVəCV and CVəCC sets where V^1 is characterised by V^1 midpoint values, is 87%-94% (mean 91%).

5.2.2.2 Range and directionality of effects

Vowel-to-vowel effects across an intervening consonant and similarly, consonant-to-vowel effects across an intervening vowel, are negligible for both F2 and F1. In the case of F2, there is a small effect of V^1 for tokens in the VCəCV set which is significant at the $p < 0.1$ level. However, with an increment in R^2 of less than 1%, the unique contribution of V^1 is minimal compared with the unique contributions made by C^1 and C^2. There is no effect of V^2. If considered alone, V^1 appears to make a small contribution to the schwa variance in the VCəCC set ($r=.199$, $p < .01$), however, this is due entirely to its relationship with C^1 ($r=.311$, $p < .01$) and through C^1, its indirect relationship with C^2. When C^1 and C^2 are partialled out, the resultant β values indicate that V^1 has no direct effect on schwa, $R=.948$, .02, $p = .35$ (V^1), .48, $p < .01$ (C^1) and .58, $p < .01$ (C^2). In the CCəCV set there is a small effect of V^2 which is just significant at $p=.04$. Addition of C^1 to the equation produces a small but significant increment in R^2 in the CVəCC and CVəCV sets (.0084 and .0156 respectively) but again, this is minimal compared with the amount of variance uniquely accounted for by each of the immediately adjacent segments.

In the case of F1, V^1 has a significant direct effect in the VCəCC set as does V^2 in the CCəCV set and both V^1 and V^2 contribute to the variance for tokens in the VCəCV set. In each case, however, the increment in R^2 is less than 1%. There is no direct effect of C^1 or V^2 in the CVəCV set and likewise no significant contribution from C^1 in the CVəCC set.

In all sets and for both F1 and F2, the immediately adjacent context segments each have a significant direct effect on schwa and together account for the major part of the variance. For F2, in two sets (CCəCV, VCəCV), the effects of preceding and following consonant are comparable. In one set only (VCəCC), the following consonant exerts a stronger influence than the preceding consonant. Otherwise, preceding context accounts for a relatively higher proportion of the variance than following context. For F1 there are two cases (CCəCC, CVəCC) in which the effects of preceding and following context are comparable. The remaining four sets display greater carryover than anticipatory coarticulation.

The greatest difference in the relative strength of carryover compared with anticipatory effects occurs in the two sets where the immediately preceding context is vocalic. This may be attributable to the fact that a relatively high proportion of the preceding vocalic context segments are /i/ tokens. These sets also include a small number of /ɔ/ tokens. In representing the upper and lower extremes of the F2 range, /i/ and /ɔ/ are likely to cause the greatest shifts in schwa F2 midpoint value. The extent of the shift is illustrated in Figure 5.29 which shows the second formant trajectories for schwa tokens in these contexts. The results of a one-way analysis of variance confirm a significant main effect of preceding vowel identity on schwa midpoint values ($F(3, 161) = 41.21$, $p < .0001$). Post hoc Tukey multiple comparison of means tests (see section 5.3) reveal significantly higher values for schwas following front vowels than following back vowels, central vowels or /u/ ($p < .001$).

5.2.2.3 Duration

For F2, there are only two cases, CCəCC and VCəCC, where addition of schwa duration to the equation results in a significant increase in R^2 (.01% and 2% respectively). In all other cases schwa duration makes no contribution to the prediction. With the exception of VCəCC, schwa duration has a significant effect on F1 in all sets. However, in each case the amount of variance it accounts for is smaller than that accounted for by either the immediately preceding or following context.

V^1 and V^2 duration have no influence on the prediction for F2. There is a small effect of V^1 duration in the VCəCC set for F1 (an increment in R^2 of 3%). There is also a small effect of V^1 duration and V^2 duration in the VCəCV set. However these do not add significantly to the overall R^2 value.

5.2.2.4 Full vowels

As in the case of schwa, the results for the full vowel subsets also showed relatively little coarticulatory effect from the next adjacent context compared with the immediately adjacent context. In view of this and in view also of the small number of tokens within some of the subsets, the analysis was repeated using data pooled across vocalic and consonantal contexts.

5.2.3 Pooled sets analysis

For the pooled data sets, tokens adjacent to silence or to the glottal fricative /h/ were excluded from the analysis. This was to ensure that differences in the degree of context-dependency between vowels were not the result of differences in the proportion of context compared with non-context segments. In addition to silence, /h/ is considered to be a non-context segment since it has a minimal coarticulatory effect on adjacent segments (Peterson & Barney, 1952; Stevens & House, 1963). The data was also screened for outliers using diagnostic statistics in the BMDP 2r regression package. The formant trajectories for all outliers identified as exerting undue force on the regression (using the modified Cook's distance statistic) were examined. In the majority of cases, the outliers represented what were considered to be legitimate, if extreme, context effects. There was a small percentage of formant tracking errors (less than 1% in the case of all vowels except /a/ and /ɑ/ which showed 1.5% errors for F2 and 2% errors for

F1 respectively). These tokens were excluded from the subsequent analyses. Results are presented in Tables 5.6:7. All R^2 values are significant at $p < .01$. With few exceptions, Beta values are significant at $p < .01$. Values which do not attain significance at $p < .01$ but are significant at $p < .05$ are marked with an asterisk (*). Those values which fail to reach significance at $p < .05$ are emboldened.

The increment in R^2 value with the addition of each predictor variable to the equation was also calculated using the BMDP 2r package. This represents the proportion of variance uniquely accounted for by each variable over and above the contribution made by the other variables already in the equation. The predictor variables were entered in the prespecified sequence: preceding context, following context, duration, according to their presumed order of causal priority. This was determined partly by the results of the initial regression analyses which, in general, show a larger effect of context than duration and a larger effect of preceding compared with following context, and partly on the basis of the results of the Tukey multiple comparison of means tests which also indicate generally greater carryover than anticipatory effects (see section 5.3.1).

5.2.3.1 Proportion of explained variance

Schwa displays the highest levels of context-dependency as indicated by the proportion of explained F1 and F2 variance. The mid-high, lax vowels /ʊ/ and /ɪ/ also display a relatively high degree of context-dependency along F2 and, in the case of /ɪ/, a relatively high degree of context-dependency along F1 also. The total proportion of explained F2 variance for /ʊ/ is comparable to that for schwa. The lower R^2 value for /ɪ/ largely reflects the influence of a specific context. The results of the data screening showed that a high percentage of the outliers identified as exerting an undue force on the regression for /ɪ/ occur in the context of a following velar nasal. Observed values for /ɪ/ in these contexts were always higher than the predicted values. Of the six formant tracking errors identified for /ɪ/ F2 trajectories, three occurred for tokens in this context.

In view of the potential for measurement errors in nasal contexts, a further regression analysis was performed. The analysis was repeated for each vowel and in each case the most extreme outliers identified by the modified Cook's distance were excluded. The results are shown in Table 5.8:9. Predictably, exclusion of the outliers results in an overall increase in R^2 value. However, the increase is most marked in the case of /ɪ/.

Since the number and nature of outliers varies considerably across vowels and since these were also considered to represent examples of legitimate if extreme contextual variation, the results of the initial analysis will form the basis of future discussion in the case of all vowels with the exception of /ɪ/. The results of the second analysis will be used in the case of /ɪ/ on account of the potential for measurement error afforded by the high proportion of following velar nasal contexts.

The R^2 value obtained for each vowel before duration is added to the equation is shown in Table 5.10. It is also illustrated graphically in Figure 5.48 (see section 5.4.2).

Table 5.10: The R^2 value obtained for each vowel before duration is added to the equation.

	ə	ɪ	ʊ	u	ɛ	ɜ	ʌ	a	i	ɒ	ɔ	ɑ
F2	.9253	.8841	.8797	.7564	.6672	.6164	.5495	.3574	.3358	.2661	.1914	.0679
F1	.7103	.7115	.2967	.4337	.2876	.1917	.2057	.1198	.3638	.2348	.0646	.0848

Following schwa and the mid-high, lax vowels /ɪ/ and /ʊ/, the tense vowel /u/ and the mid-low, lax, front vowel /ɛ/, the long, central vowel /ɜ/ and the mid-low, lax, back vowel /ʌ/, in descending order, show the highest level of F2 context dependency. Among the remaining vowels, the back vowels /ɑ/, /ɔ/ and /ɒ/ show the least context-dependency along F2 followed, in ascending order, by the tense, front vowel /i/ and the long, lax vowel /a/.

The ranking of vowels from least to most context-dependent along F1 is similar insofar as the long vowels /ɑ/, /ɔ/ and /a/ show the least context-dependency while schwa and /ɪ/ together with /u/ show the highest context-dependency. The mid-high, lax vowel /ʊ/, however, shows relatively low context-dependency along F1 while the tense, front vowel /i/ occupies a relatively low position in the ranking of vowels from least to most context-dependent. With the exception of /i/, all the vowels show lower context-dependency along F1 than along F2.

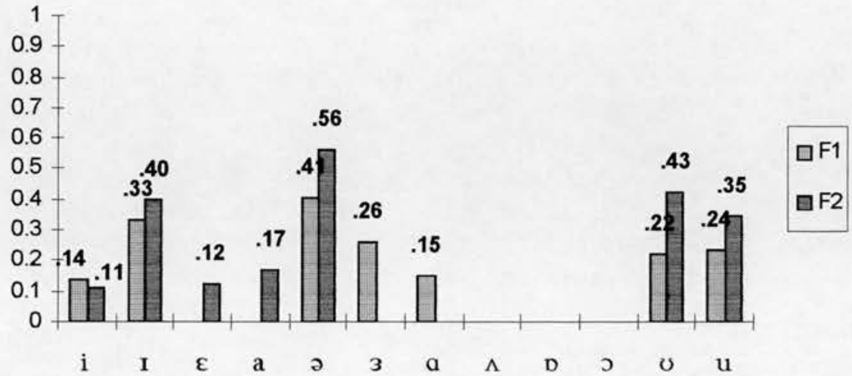
The extremely high correlation between vowel F2 midpoint values and context for /ə/, /ɪ/ and /ʊ/ is illustrated graphically in Figures 5.30:32. For comparison, Figures 5.33:34 show the scatterplots for /i/ and /ɔ/, two of the more stable vowels. The differences in degree of context-dependency across vowels is considered in more detail in section 5.4.2.

5.2.3.2 Directionality of effects

Overall, the results indicate stronger carryover than anticipatory coarticulation. Following context makes the largest contribution to the F2 prediction for /ɛ/. For all other vowels, preceding context accounts for more of the F2 variance than following context although the difference is marginal in the case of /a/, /ʊ/ and schwa. In the case of F1, there is a stronger anticipatory than carryover effect for /ɜ/, /ɛ/, /ʉ/ and /ʊ/. All other vowels, including schwa, show a stronger carryover effect. For /ɔ/, there appears to be no anticipatory coarticulation along either F1 or F2. Similarly, following context has no unique F2 variance in /ɑ/ and accounts for only a small proportion of the F2 variance in /ɒ/.

In order to assess the interaction of context effects through a medial vowel, the correlation between preceding and following context was also calculated. This is represented graphically in Figure 5.35. Only those values which are significant at $p < .05$ are shown.

Figure 5.35: Correlation between F1/F2 onset values and F1/F2 offset values



Schwa, /ʊ/ and /ɪ/ show the greatest correlation between preceding and following context, indicating a moderately strong interaction of context effects through the medial vowel. There is no interaction of effects for the back vowels /ʌ, ɒ, ɔ/ along either F1 or F2. In general, there is less interaction of effects for F1 than for F2.

5.2.3.3 Duration

As the Beta values given in Table 5.6 demonstrate, duration makes little or no contribution to the F2 prediction in the case of /ə, a, ɜ, u/ and a relatively small contribution to the prediction for /ɪ, ɛ, ʊ, ʌ/. However, in the case of /i, ɑ, ɔ, ɒ/ it accounts for a higher proportion of F2 variance than following context. For F1, duration uniquely accounts for the highest proportion of variance in the case of /a, ɜ, ʌ/ and for a higher proportion of F1 variance than following context in the case of /ɑ/. It also makes a greater contribution to the F1 prediction than preceding context in the case of /ɛ/.

The proportion of variance in midpoint value for each vowel that is uniquely accounted for by duration is shown in Table 5.11.

Table 5.11: Increment in R^2 value with the addition of duration to the regression equation for each vowel.

	i	ɪ	ɛ	a	ə	ɜ	ɑ	ʌ	ɒ	ɔ	ʊ	u
F2	19%	2%	3%	1%	0	0	5%	3%	18%	19%	2%	0
F1	3%	0	13%	24%	3%	19%	5%	23%	10%	0	5%	2%

As this table demonstrates, duration accounts for a higher proportion of the variance in F2 midpoint value in the case /i/, /ɔ/ and /ɒ/ which represent the extreme upper and lower limits in F2 value than in the case of the more central vowels. It also accounts for a higher proportion of the variance in F1 midpoint value for the low and mid-low vowels /a/, /ɜ/, /ʌ/ and /ɒ/ than for the other vowels.

5.3 Degree of differentiation in vowel midpoint value: analyses of variance

In order to assess the degree of differentiation in vowel midpoint value as a function of preceding (C^1) or following (C^2) consonantal place of articulation, a series of one-way analyses of variance were performed. (Variation in midpoint value as a function of adjacent vocalic context was not considered in the analyses of variance owing to the relatively small number of vowel tokens occurring immediately adjacent to another vowel.) Pairwise differences were assessed using the post hoc Tukey multiple comparison of means test. Because of the high number of simultaneous tests, significance levels for this were Bonferonni adjusted. In many cases, where unequal variances obtained between cells, as indicated by Levene's test for equal variability (see BMDP, p. 192), the Brown-Forsythe separate variance procedure was used in place of the standard pooled analysis of variance. Although this test involves a loss of degrees of freedom, it is more robust against unequal group variances. Separate variance t-tests were also used where appropriate in the pairwise multiple comparisons. Unequal sample size between cells are controlled for automatically in the BMDP 7D package using the Tukey-Cramer adjustment.

Two-way analyses of variance were also performed to ensure that potential differences in results between vowels are not due to differences in the extent to which C^1 and C^2 place effects or C^1/C^2 place and manner effects interact. Due to the limited number of tokens occurring within some $C^1$and C^2 combined contexts, for the C^1xC^2 place analysis, the eight place categories were collapsed into three levels: High, Mid and Low. Classification as 'High', 'Mid' or 'Low' was performed on the basis of the relative frequency value of F2 at vowel onset and offset as a function of $C^1$and C^2 place of articulation. Figures 5.36:39 shows the relative ranking of place categories on the frequency scale at onset and offset for each vowel. In general, palatals and alveolars are characterised by relatively high onset/offset values compared with labials, apicals and labio-velar /w/. Velars are characterised by relatively high values for all vowels except the rounded vowels /ɒ/, /ɔ/, /u/ and /ʊ/ where they are characterised by relatively low values, particularly at vowel offset. The relative ranking of dentals and labio-dentals also varies as a function of temporal position. At vowel onset, dentals are generally characterised by higher F2 values than labio-dentals. At vowel offset, there is less difference in frequency value between the two place categories in the case of /ə/, /i/, /a/ and /ɒ/ owing to relatively higher offset values for labio-dentals. Lower absolute values and a relatively lower ranking of apicals at offset than at onset largely reflects the allophonic variation between 'clear' /l/ which occurs in pre-vocalic position, and 'dark' or velarised /ɫ/, which occurs post-vocalically.

The place x manner two-way analysis of variance was performed for labials and alveolars by including labio-dentals as labial fricatives. The same measurement and classification procedures were used for both F1 and F2.

5.3.1 F2

The results for F2 are considered first. These are presented in Tables 5.12:5.17. Table 5.12 gives the results of the one-way analysis of variance while Table 5.13:14 shows the number of pairwise differences that were significant ($p < .05$) for each vowel in the

Tukey multiple comparison of means test. A significant difference is indicated by the presence of the appropriate vowel symbol within a given cell. Thus, for example, a significant difference in F2 for tokens in C^1=palatal compared with C^1=velar contexts obtains for /ə/, /i/, /ɪ/ and /ɛ/. The vowel mean midpoint values for each C^1/C^2 place of articulation are also plotted in Figures 5.40:5.43. Results for the two-way analyses of variance are given in Tables 5.15:17.

As expected, schwa displays the highest proportion of significant pairwise differences between place categories at both the carryover and anticipatory level. For midpoint values grouped according to C^1 place of articulation, most of the overlap which occurs is between place categories characterised by low F2 values. There is no significant place effect for labio-dentals compared with labials and labio-velar /w/, for labials compared with apicals and /w/ or for apicals compared with /w/. Overlap also occurs between schwa tokens in preceding velar, alveolar and dental contexts. However, F2 is significantly higher for schwa tokens in the context of a preceding palatal, velar, alveolar or dental than all other contexts ($p < .001$). Values are also significantly lower following /w/ than following /l, r/ ($p < .05$). Schwa midpoint values grouped according to C^2 place of articulation show greater overall separability. The only overlap which occurs is between tokens in following dental and labio-dental contexts. All other pairwise differences are significant at $p < .001$.

The high degree of separability in /ə/ midpoint value as a function of C^1 and C^2 place of articulation is particularly noteworthy given the interaction of context effects for /ə/ and the asymmetry of the majority of combined C^1 and C^2 contexts. As Figures 5.1:5.5 demonstrate, the ranking of mean midpoint values for each C^1 place category is, to a large extent, conditioned by the place of articulation of the following consonant such that, for example, in Figure 5.2a, midpoint values decrease in the following order Vel-Pal > Vel-Alv > Vel-Labd > Vel-Lab > Vel-Api. The C^2 influence at the /ə/ midpoint and onset is most evident in the case of preceding velar, dental and labio-dental contexts. It is least evident in the case of preceding labial and apical contexts. The ranking of mean midpoint values within each C^2 place category is similarly influenced by the place of articulation of the preceding context (see Figure A 2, Appendix A). The interaction of C^1 and C^2 place effects at the schwa midpoint is confirmed statistically in the two-way analysis of variance.

The F2 midpoint values for /ɪ/ show a lesser degree of separability as a function of preceding consonantal place of articulation than schwa midpoint values, with overlap occurring between tokens following palatals, alveolars, dentals, labials and labio-dentals. However, tokens following velars, apicals and /w/ are clearly differentiated from tokens in all other contexts, attaining maximum values in the velar context and minimum values in the apical and labio-velar contexts (see Table A 1.b, Appendix A for the mean values). As in the case of schwa, greater overall separability in /ɪ/ midpoint values occurs as a function of C^2 place of articulation. Overlap occurs between tokens preceding alveolars, labials and labio-dentals, between tokens in dental and apical contexts and between palatals and labio-velars. Overlap between palatals and labio-velars reflects the small number (five) and extremely high variability of tokens preceding /w/. With these exceptions, the differences in midpoint value for tokens in all other contexts attain statistical significance: for palatals compared with alveolars and labio-dentals at $p < .01$ and for all other pairwise comparisons at $p < .001$.

In the case of /ɛ/, midpoint values are significantly higher in the context of a preceding velar than in any other context other than a preceding dental. Overlap in this instance is due to the small number (six) and high variance of tokens in the dental context. As in the case of /ɪ/ and /ə/, the lowest values occur for tokens following /w/. Values are also significantly lower in the context of a preceding labio-dental than all contexts except for /w/. The influence of C^2 place of articulation is more constrained and almost entirely attributable to the lowering effect of velarised /ɫ/.

At the carryover level, F2 midpoint values for /ʊ/ vary significantly as a function of high compared with low locus consonants, being higher in the context of a preceding palatal, alveolar or velar than in other contexts. The lowest values occur following labio-dentals and apicals. At the anticipatory level, however, there is no apparent differentiation in value as a function of consonantal place of articulation.

Results for /ʊ/ reveal comparatively little separability in F2 midpoint value as a function of either preceding or following consonantal place of articulation. Values are significantly higher following palatals than following velars ($p < .1$) or apicals ($p < .01$). Otherwise there is no differentiation between place categories at the carryover level. As in the case of /ɛ/, the main effect of C^2 observed for /ʊ/, reflects the influence of a following dark /ɫ/. Midpoint values for /ʊ/ are significantly lower preceding apicals than in all other contexts.

The lower overall context-dependency displayed by the mid-low, back vowel /ʌ/ compared with the corresponding mid-low, front vowel /ɛ/, is largely attributable to a lesser degree of anticipatory coarticulation. This, in turn, may be attributed to the fact that velarised /ɫ/ which accounts for the anticipatory effect on /ɛ/ midpoint values, is articulatorily and acoustically compatible with back vowels and therefore does not exert the same coarticulatory influence (Recasens, 1991). By the same token, a strong anticipatory effect for /ʌ/ might be expected in the context of a following palatal or alveolar. The absence of such an effect in the present data is due, in part, to the small number of following palatal contexts (five tokens). It is noteworthy, however, that there is no significant effect of a following alveolar. A possible explanation lies in the interaction of context effects whereby the raising effect of a following alveolar is cancelled out by the lowering effect of a preceding low locus consonant. The mean midpoint value for /ʌ/ in C^2=alveolar contexts is lower when the opposing C^1 context is a low locus consonant than when it is a high locus consonant (see Table A 1.g, Appendix A). However, the results of the two-way analysis of variance do not show a significant $C^1 \times C^2$ interaction. Furthermore, corresponding variation in midpoint value for /ʌ/ in C^1=alveolar contexts as a function of C^2 place of articulation is not apparent also indicating greater robustness for alveolars in C^1 position.

At the carryover level, /ʌ/ midpoint values are significantly higher following palatals ($p < .01$), velars ($p < .001$) and alveolars ($p < .001$) than following labials or /w/. There is also a significant difference in midpoint value for tokens in preceding alveolar contexts compared with tokens in preceding labio-dental ($p < .01$) or apical ($p < .001$) contexts. Values are also significantly lower in the context of a preceding /w/ than a preceding dental ($p < .05$) or apical ($p < .001$).

The highest midpoint values for /ɒ/ also occur in the context of a preceding alveolar or velar and the lowest values, in the context of preceding labio-dentals, labials and apicals. The pairwise differences between groups are significant in the case of the alveolar compared with the labio-dental, labial and apical categories ($p < .001$) and in the case of the velar compared with the labio-dental ($p < .05$) and labial ($p < .01$) categories. For /ɔ/, significant pairwise differences in midpoint value are obtained for tokens in the context of a preceding velar compared with tokens following labio-dentals ($p < .05$), labials ($p < .01$) and /w/ ($p < .01$), also in the context of preceding alveolars compared with labio-dentals ($p < .01$) or labials and labio-velar /w/ ($p < .001$). Significant pairwise differences are also obtained for preceding labio-dental ($p < .05$) and labial ($p < .01$) contexts compared with apical contexts and for apical compared with labio-velar contexts ($p < .01$). Higher values are obtained in the velar, alveolar and apical contexts. There is no significant main effect of following consonantal place of articulation for either vowel. As for /ʌ/, results of the two-way analysis of variance indicate that the lack of a C^2 effect cannot be attributed to a $C^1 \times C^2$ interaction.

As in the case of /ɛ/, the anticipatory effect obtained for /i/ and /a/ is largely due to the influence of a following velarised /ɫ/. In both cases, carryover coarticulatory effects are more evenly spread across contexts. For /i/, the highest values predictably occur in the context of a preceding velar or alveolar and the lowest values, following /w/. Significantly lower values are also obtained in the context of a preceding apical than in the context of a preceding labial ($p < .01$). For the low, front vowel /a/, significant differences are obtained for tokens in the context of a preceding labio-dental ($p < .00$) compared with tokens following velars and alveolars ($p < .01$) and between tokens in the context of preceding velars and apicals ($p < .01$), with lower values occurring in the labio-dental and apical contexts. Significantly lower values are also obtained in a preceding labio-dental context than in either a preceding labial or apical context ($p < .01$).

The central vowel /ɜ/ also displays significant differences in midpoint value for tokens following palatals, velars and alveolars compared with tokens following labials ($p < .001$), labio-dentals ($p < .01$) and /w/ ($p < .05$). However, as in the case of the back vowels, midpoint values show no differentiation across palatal, velar and alveolar contexts. Overlap also occurs between dental, labio-dental and labial contexts. Values are significantly lower following /w/ than in all other contexts except following labials and labio-dentals. Again, there are a greater number of significant pairwise differences at the carryover than at the anticipatory level. The only significant difference in midpoint value that occurs as a function of following consonantal place of articulation is for tokens in velar and labial contexts compared with tokens in alveolar contexts, with higher values being obtained in the alveolar contexts ($p < .001$ and $p < .05$ respectively).

For the low, back vowel /ɑ/, there is no main effect of either C^1 or C^2 place of articulation and no $C^1 x\ C^2$ interaction. The stability of this vowel across different contexts is illustrated graphically in Figure 5.45 which shows the /ɑ/ mean second formant trajectories plotted as a function of both C^1 and C^2 place of articulation.

In sum, F2 midpoint values for front vowels tend to be higher following velars than in all other contexts. Relatively high values for /i, ɪ, a/ also occur following alveolars. In the case of /ɛ/, values are higher in the context of a preceding palatal than a preceding alveolar. This is also true in the case of /ʊ/, /ɯ/ and /ɜ/. For the back vowels, higher values generally occur following alveolars, although for /ɑ/ and /ɔ/, F2 is higher following apicals. Relatively high values also occur for /ɒ/ and /ɔ/ in the context of a preceding velar or palatal. The lowering effect of the labio-velar glide /w/ is highly systematic and consistent across all vowels although, in terms of the amount of shift in F2, is predictably of a lesser extent for the back vowels than for the front vowels. There is generally less differentiation in vowel midpoint value at the anticipatory than at the carryover level. In the case of the front vowels and /ɯ/, /ʊ/ and /ɜ/, a significant C^2 effect largely reflects the lowering influence of a following velarised /ɫ/. The two exceptions to this are /ə/ and /ɪ/ which show a higher degree of separability in midpoint value as a function of C^2 than of C^1 place of articulation. Schwa, closely followed by /ɪ/, shows the greatest differentiation in F2 midpoint value overall.

The results of the two-way analysis of variance and pairwise comparisons for place x manner effects generally reveal lower F2 values for vowels in fricative compared with stop and nasal contexts. For example, lower F2 midpoint values for /ɪ/ and /ɛ/ are obtained in the context of a preceding alveolar fricative than in either the corresponding stop or nasal contexts. In the case of /ɪ/, both pairwise comparisons are significant at $p < .001$. In the case of /ɛ/, there is a greater difference between the fricative and nasal ($p < .001$) than between the fricative and stop contexts ($p < .05$). /ɛ/ also shows a significant difference in value for tokens in the context of preceding /f, v/ compared with tokens following /p, b/ ($p < .01$) or /m/ ($p < .001$). Schwa also displays lower values following /s, z/ than /t, d/ ($p < .01$) and lower values following /f, v/ than following /m/ ($p < .05$). Similarly, lower values are obtained for /a/ in the context of a preceding alveolar or labial fricative than in the corresponding stop contexts ($p < .05$). There is also a significant difference in value for /a/ tokens following /f, v/ compared with tokens following /m/ ($p < .05$). In the case of /i/, lower F2 midpoint values are obtained for tokens in the context of a preceding /t, d/ compared with tokens following /n/ and lower values for tokens following /p, b/ and /f, v/ than for tokens following /m/ ($p < .001$ and $p < .01$ respectively).

/ɪ/ and /a/ are the only vowels to show an interaction of C^1 place x manner effects. In both cases, this reflects higher values for tokens in alveolar contexts compared with in labial contexts except when the alveolar is a fricative or the labial is a nasal. The back vowels and /ɜ/ do not show any manner effects within either place category.

Manner effects are less pervasive at the anticipatory level. Lower F2 midpoint values are obtained for /i/ tokens preceding /s, z/ than preceding /p, b/ ($p < .01$) and for /a/ tokens preceding alveolar fricatives and stops than preceding /n/ ($p < .001$). Otherwise there are no significant manner effects. /u/ shows a significant interaction of effects such that lower mean values are obtained in the context of a following alveolar than a following labial unless the alveolar or labial is a fricative. The only difference to attain significance, however, is between tokens in the alveolar and labial fricative contexts ($p < .05$).

In general, the observed pattern of consonantal place effects is consistent with effects reported in the literature (Stevens & House, 1963; Stevens, House & Paul, 1966). The ranking of consonantal place categories on the high-low F2 continuum are also consistent with the acoustic resonator theory (Stevens & House, 1956) which predicts increasingly lower F2 locus values as consonantal place of articulation moves forward in the vocal tract. The exception to this occurs in the case of velars in the context of back rounded vowels where they are characterised by relatively low onset/offset values. In the case of the labio-velar glide /w/, lip-rounding serves to increase the length of the vocal tract, thereby lowering F2 relative to the other labial consonants (Lehiste, 1964; Mack & Blumstein, 1983). Despite considerable variation in absolute onset/offset value as a function of vowel identity, the high-low rank-ordering of consonantal place categories is relatively stable across front/back vowel classes and individual vowels (see Figures 5.37:5.39).

The lower values observed for front vowels and for schwa in both alveolar and labial fricative compared with in the corresponding stop or nasal contexts is also consistent with observations in the literature (Stevens & House, 1963; Recasens, 1991; Farnetani et al., 1993). The greater coarticulatory influence of fricatives in these studies is attributed to greater mechanical constraints on the movements of the tongue and jaw during their production. According to Stevens & House, the slower and more precise movements involved in achieving an appropriate fricative constriction compared with complete closure for stops, results in less time being available for the vowel and thus greater undershoot. In gestural terms, the production requirements of the fricative take precedence over those of the vowel (Recasens, 1991; Farnetani et al., 1993).

5.3.2 F1

One-way analyses of variance were performed for F1 in which the grouping variable was either C^1/C^2 place of articulation or C^1/C^2 manner of articulation. A two-way analysis of variance was also performed to test for place x manner effects. As in section 5.3.1, this examined for differences in vowel midpoint value as a function of manner of articulation in alveolar or labial (including labio-dental) contexts. Results of the analyses of variance are presented in Tables 5.18:5.22. Vowel mean midpoint values are plotted in Figure 5.44.

As Table 5.18 demonstrates, place effects for F1 are predictably less extensive than for F2. In addition to schwa, only the front vowels /ɪ/, /i/, /ɛ/ and /a/ show a significant C^1 main effect. In the case of /i/, this reflects a difference in F1 midpoint frequency for tokens in velar compared with dental, labio-velar and labial contexts ($p < .05$). Values are highest following dentals and labio-velar /w/ and lowest, following velars. In the case of /ɪ/, the C^1 effect is accounted for solely by the raising influence of a preceding apical. Values are significantly higher in this context than following dentals ($p < .01$) or following palatals, velars and alveolars ($p < .001$). F1 midpoint values are significantly lower for /ɛ/ in the context of a preceding velar or labio-velar than a preceding alveolar ($p < .01$). For the low, front vowel /a/, F1 midpoint values are significantly higher in the context of a preceding labio-dental ($p < .05$), labial or apical ($p < .01$) than in the context of a preceding dental. Schwa displays the highest degree of separability. The lowest values occur following /w/ and palatals and the highest values, following apicals and labials. Overlap occurs between values for tokens in preceding palatal compared with labio-velar contexts, velar compared with alveolar, dental, labio-dental and labial contexts, alveolar compared with dental contexts and labial compared with labio-dental and apical contexts. All other pairwise differences attain significance at $p < .01$ (or $p < .05$ in the case of the Pal-Alv and Pal-Den comparisons).

C^2 place of articulation has a wider influence than C^1 place of articulation with the back vowels /ɒ/, /ʌ/ and /ɯ/ also showing a significant main effect in addition to schwa, /ɪ/, /ɛ/ and /a/. Note, however, that there is no significant C^2 place effect for /i/. In the case of /ʌ/ and /ɯ/, the C^2 effect reflects the influence of a following apical. This is to raise the F1 midpoint values for /ʌ/ and to lower them for /ɯ/. A following apical also serves to lower F1 in the case of /a/. The highest values for /a/ occur in the context of a following labial. In the case of /ɛ/, F1 is significantly higher in value for tokens preceding labials compared with palatals ($p < .001$), velars ($p < .05$) and alveolars ($p < .01$). /ɒ/ midpoint values are significantly higher in the context of a following velar than a following alveolar ($p < .05$), otherwise there are no significant pairwise differences. Schwa and /ɪ/ show the greatest differentiation in F1 midpoint value. For /ɪ/, values are highest in apical and labial contexts while for /ə/, they are highest in apical and labial/labio-dental contexts. Relatively low values for /ɪ/ occur preceding dentals and labio-velar /w/ and for schwa, preceding velars and palatals. The difference in /ə/ F1 midpoint value is significant for tokens preceding velars and palatals compared with alveolars, labials, labio-dentals ($p < .01$) and apicals ($p < .001$) and for tokens preceding apicals compared with alveolars, dentals, labio-dentals and /w/ ($p < .001$).

A significant main effect of C^1 manner of articulation is evident for all vowels except /ɜ/, /ɑ/, /ɒ/ and /ɯ/. The mid vowels /ɒ/ and /ɜ/ also fail to show a significant C^2 main effect. There is also no significant C^2 effect for /ʊ/. At the carryover level, the F1 values for /i/ are significantly lower following stops than in all other contexts and significantly higher following nasals than all other contexts apart from glides. For /ɛ/ there is a significant difference between values for tokens following nasals compared with glides ($p < .01$), the lower values occurring in the glide context. For /a/, values are significantly higher in the context of a preceding liquid than following fricatives (p < .05). Significantly higher values occur following nasals compared with fricatives for /ɔ/ (p < .05) and for nasals compared with fricatives (p < .01) and stops (p < .001) in the case of /ʌ/. As with place of articulation, schwa and /ɪ/ show the greatest degree of differentiation in F1 midpoint value as a function of different manner categories. In the case of schwa, overlap occurs between tokens in stop and fricative contexts and between tokens following nasals and liquids, otherwise all pairwise differences are significant (p < .01). As for the other vowels, higher values generally obtain in the nasal and liquid contexts and lower values in the stop and fricative contexts. The lowest values occur for tokens following glides. /ɪ/ also displays significantly lower F1 values following stops than in any other context. Values are also significantly higher in preceding nasal and liquid contexts than in all other contexts.

A similar pattern is also evident at the anticipatory level with generally higher values occurring in nasal and liquid contexts and lower values in stop and fricative contexts. F1 values for /i/ tend to be lower preceding stops than preceding nasals (p < .001) and liquids (p < .01). The difference between group means is also significant for fricative compared with nasal contexts (p < .001). Following nasal contexts also display higher values than following stop contexts (p < .001) in the case of /ɜ/, than following stop and fricative contexts (p < .01) in the case of /ɯ/, than following liquids (p < .05) in the case of /a/ and than following stop and liquid contexts (p < .01) in the case of /ɔ/. Significantly higher values for nasal contexts than for stop or liquid contexts (p < .001) are also evident in the case of /ʌ/. For /ɛ/, F1 values are significantly higher for tokens in both preceding nasal (p < .001) and liquid (p < .01) contexts than in preceding stop contexts. The lowest F1 values for /ɑ/ occur following glides (p < .001) and also tend to be lower preceding stops compared with nasals (p < .01) and fricatives (p < .05). Overlap occurs for schwa tokens in the context of following stops, fricatives and glides and in the context of following nasals and liquids. All other pairwise differences are significant at p < .001. As before, higher values obtain preceding liquids and nasals and lower values preceding glides, stops and fricatives. /ɪ/ is also characterised by higher F1 values in the context of a following nasal or liquid than a following stop, fricative or glide. The difference in value for tokens in the liquid and nasal compared with the glide context are significant at $p < .05$. All other pairwise differences are significant at $p < .001$.

The results of the two-way analysis of variance for C^1 place x manner for alveolar and labial consonants (see Table 5.21) confirm that manner of articulation is more influential than place of articulation. In accordance with the C^1 one-way analysis of variance results, the main C^1 manner effect is largely accounted for by a difference in F1 value for tokens in nasal compared with stop contexts. For schwa and /ɪ/ the distinction applies across both place categories. For both vowels, F1 is significantly higher following /n/ than following either the stops /t, d/ or the fricatives /s, z/ ($p < .001$). Within the labial category, schwa values are significantly higher following /m/ than following the labio-dental fricatives /f, v/ ($p < .05$). For /ɪ/, significantly higher values obtain in the context of a preceding /m/ compared with the corresponding stop context ($p < .001$) but there is no distinction between the nasal and fricative categories. Results for /u/ show the same pattern as for /ɪ/ attaining significance at $p < .05$. In the case of /i/, the distinction only applies in the case of alveolars. Values are significantly higher in alveolar nasal contexts than in either alveolar fricative ($p < .05$) or stop contexts ($p < .001$). F1 is also significantly higher following /m/ than /t, d/ ($p < .05$) although there is no distinction between /m/, /p, b/ or /f, v/. For /ʌ/ and /ɔ/, the manner effect is less extensive again, reflecting a distinction only between /n/ ($p < .01$) and /m/ ($p < .05$) and the labial stops in the case of /ʌ/ and, in the case of /ɔ/, a distinction between /m/ and the alveolar fricatives ($p < .05$).

Results show a significant place x manner interaction only in the case of schwa and /ɪ/. For schwa, this reflects lower values for tokens in alveolar compared with labial contexts except when the alveolar is nasal. In the case of /ɪ/, lower values obtain for alveolar compared with labial contexts except when the alveolar is nasal or when the labial is a stop.

Effects at the anticipatory level show a similar pattern to those at the carryover level, although there are some differences. In general, C^2 manner effects are more pervasive than C^1 manner effects. In addition to /ə, ɪ, i, ʌ, ɔ, u/, significant main effects are also apparent for /ɑ/, /ɛ/ and /ɜ/. For the most part, there are also a greater number of significant pairwise differences as a function of C^2 compared with C^1 manner of articulation. For example, in the case of /ɪ/, there is also a difference in F1 for tokens preceding alveolar stop compared with fricative contexts, lower values being obtained in the fricative contexts ($p < .01$). Similarly, there is also a distinction between labial stops and labio-dental fricatives for /u/ at the anticipatory level ($p < .05$) which is not evident at the carryover level. However, in this case, lower values are obtained in the stop context. A significant interaction of effects for /u/ also results in higher F1 values for labial compared with alveolar contexts except when the labial is a stop. Unlike at the carryover level, there is no significant manner effect for following alveolars.

In the case of /i/, there is an additional manner effect for labials at the anticipatory level, such that higher values obtain for tokens preceding /m/ than for tokens preceding /p, b/ ($p < .05$). However, unlike at the carryover level, F1 is not significantly higher for tokens preceding /m/ compared with tokens preceding alveolar fricatives and stops. For /ʌ/, there is an additional distinction between alveolar nasal and alveolar stop contexts ($p < .01$), with higher values occurring in the nasal contexts. However, the distinction between preceding alveolar nasal and labial stop contexts no longer applies. F1 values for /ɛ/ are significantly higher preceding /n/ than preceding the alveolar stops ($p < .01$). There is also a significant place effect: F1 is higher preceding /m/ than preceding /n/ or /s, z/ ($p < .05$) and /t, d/ ($p < .001$). In the case of schwa, there is a reduction in the number of pairwise differences that are significant at the anticipatory level compared with at the carryover level. As before, there are significant manner effects within the alveolar category and in the same direction, however, there are no manner effects within the labial category.

In comparison with F2, there is relatively little information in the literature with respect to consonant-to-vowel coarticulation for F1. There is also less consistency in the observed effects across studies. This may be attributable, in part, to between speaker variability with respect to the way in which tongue height is controlled. While for some speakers, tongue height is dependent on jaw position, for other speakers, tongue and jaw are to some extent moved independently of each other. There is also evidence that the extent to which tongue and jaw interact varies depending on context (Lindau & Ladefoged, 1990). Thus, less correspondence between studies with respect to F1 contextual variation than in the case of F2 variation may reflect greater between-speaker differences in articulatory strategy for tongue height relative to front/backness.

With respect to the place effects observed in the present study, the generally higher F1 values for vowels in labial contexts is consistent with observations made by Lindau & Ladefoged (1990). They report that three of the four speakers in their study showed a lower jaw position for bilabials than for alveolars. They also note that this contrasts with the observations of Keating et al., (1990) who report a higher average jaw position for vowels in the context of /b/ than in all other consonantal contexts. Lindau & Ladefoged explain their results in terms of the rapid lip movement towards bilabial closure that was also observed for their speakers. In the bilabial context, the lower lip was found to travel faster and further than in the alveolar context. The implication is that a lower jaw position in the bilabial context is tolerated because consonantal closure can be achieved by rapid lip movement.

Lower F1 values observed for vowels in fricative contexts reflects the higher jaw position necessary to achieve the appropriate placement of the upper and lower teeth to produce high frequency friction. Generally higher values for tokens in liquid contexts also reflects a comparatively low tongue-body and/or jaw configuration for vowels in this context (Farnetani et al., 1993). Higher values obtained for vowels in nasal contexts, however, may be attributed to the effects of nasalisation. In their study of vowel nasalisation using electrical analogs of the oral and nasal tracts, House & Stevens (1956) report an increase in the frequency of F1 for all vowels with nasal coupling. In the present study, a relatively small proportion of the vowels which occur in nasal contexts were transcribed as being nasalised. However, given that nasalisation is not used contrastively for vowels in English, this may reflect a lack of refinement in the labeller's perception of nasality. It is also the case that vowels in connected speech are likely to show varying degrees of velopharyngeal opening which may not always be sufficient to cue the perception of nasality.

5.4 Discussion

In the following discussion, the magnitude of coarticulatory effects and the patterns of variability displayed by different vowels are evaluated with reference to the question of inherent vowel variability. The results are considered first for schwa and then for the full vowels.

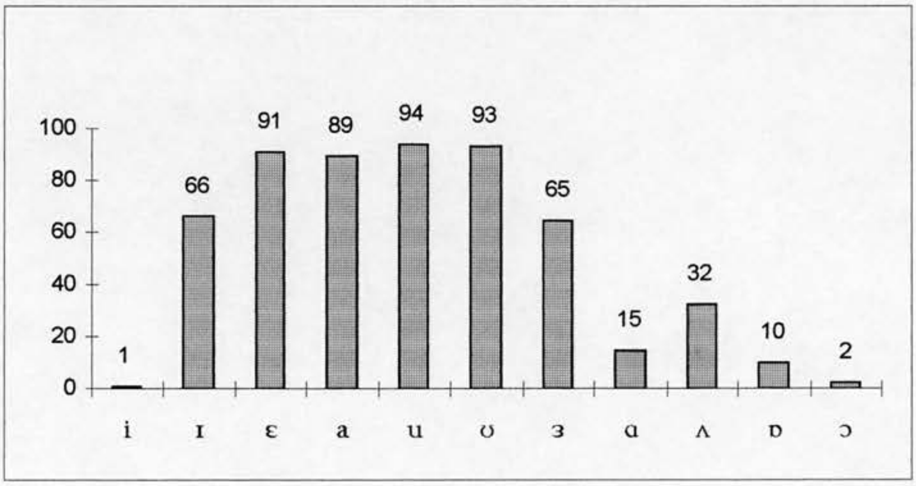
5.4.1 Schwa

5.4.1.1 F2

The results of the regression analysis provide strong support for the theory that schwa is unspecified for tongue position. Given that 100% prediction accuracy is unlikely owing to the inevitable presence of some random variability, 92% explained F2 variance for schwa tokens in the pooled data set (94% for the CCəCV set) arguably denotes a near maximal level of context-dependency. The linearity of the majority of schwa F2 mean trajectories and the wide spread in schwa F2 midpoint value also indicate interpolation through, although the overall range in schwa F2 midpoint value, as shown in section 5.1, is subject to some constraint.

Browman & Goldstein (1992) report that in their articulatory data, schwa shows greater variability in tongue position than any other vowel but that the range of this variability does not extend across the entire vowel space. The present acoustic data accords with this observation. As demonstrated in section 5.1, the F2 mean midpoint values for schwa do not extend into the extreme upper or lower regions in F2 space. The highest mean value for schwa obtained in the Pal-Vel context (1786 Hz) is not as high as the corresponding mean values obtained for /i/ and /ɪ/ (2106 Hz and 1850 Hz respectively) (see Table A 1.a:b, Appendix A). Similarly, the lowest mean value for schwa obtained in the Labd-Labv context (958 Hz) is not as low as the values obtained for /ɔ/ and /ɒ/ in the near equivalent contexts Labd-Lab and Lab-Api (746 Hz and 878 Hz respectively). In order to determine how far these figures reflect what occurs at the individual token level, the percentage of overlap along F2 between /ə/ and /i/ tokens and between /ə/ and /ɔ/ tokens was calculated. A cut-off of 5% was used to exclude outliers and to maintain consistency with the vowel ellipses plotted in F1-F2 space and shown in section 6.2. For comparison, the percentage of overlap along F2 between schwa and each of the other vowels was also calculated. This is shown in Figure 5.46.

Figure 5.46: Percentage of overlap along F2 between schwa and the full vowels.



The minimal degree of overlap between /ə/ and /i/ and between /ə/ and /ɔ/ tokens and the relatively small degree of overlap between schwa and the other back vowels (excepting /u/ and /ʊ/) confirms that there are constraints on the spread in schwa F2 midpoint value relative to the maximum range possible in F2. These constraints, however, do not reflect an inherent front/back specification but are rather reflective of contextual limitations. The only consonants which are characterised by F2 values approaching the target F2 values for /i/ or /ɔ/ are those with primary or secondary palatalised or velarised articulations. Thus, the only context in which /ə/ is likely to approach the values characteristic of /i/ is in the context of two palatals, between two /i/ vowels or between /i/ and a following velar which has assimilated the /i/ high F2 value. In the present data, there are no examples of schwa in the context of a preceding and following palatal or /i/ vowel. However, there are five examples of schwa in the context of a preceding /i/ and a following velar. The mean midpoint value for these tokens at 1955 Hz is considerably higher (169 Hz) than the corresponding mean obtained for tokens in the Pal-Vel context.

In the case of /ɔ/, similarly low values for schwa are only likely to occur in the context of /ɔ/ (or /u/ for speakers who produce this with a full back quality) and /w/ or /ɫ/. In the present data, there is one example of schwa occurring in a symmetrical labio-velar context and five tokens which occur in the context of a preceding /ɔ/ and following /ɫ/. In these contexts schwa displays a midpoint value and a mean midpoint value of 922 Hz and 1065 Hz respectively. Thus, it appears that, given the appropriate contexts, schwa F2 midpoint values do approach the upper and lower limits in F2 value, although they do not attain the most extreme values characteristic of the majority of /i/ and /ɔ/ tokens.

The temporal extent and directionality of effects for /ə/ also indicate phonetic transparency. In line with Keating's (1988) predictions for unspecified segments, coarticulatory effects extend throughout schwa in both directions such that there is a moderately strong C^2 influence at schwa onset and a moderately strong C^1 influence at schwa offset. This is reflected in the correlation between onset and offset value ($r = .56$ for F2 and $r = .41$ for F1). In addition to a greater interaction of context effects, and also in accord with Keating's predictions, there is also a greater parity in the level of carryover compared with anticipatory coarticulation for schwa (and /ɪ/) than for the full vowels.

In terms of the overall magnitude of carryover compared with anticipatory effects, the results of the regression analyses indicate a marginally stronger carryover effect while the analyses of variance show a stronger anticipatory influence. The discrepancy in results may be attributed to the fact that only adjacent consonantal context is considered in the analyses of variance. The relatively small number of immediately adjacent vocalic contexts prohibited comparison of individual vocalic place of articulation. In the regression analysis all tokens are included because no distinction is made between individual contexts. However, given that amount of coarticulation varies depending on the specific combination of vowel and consonant gestures (Stevens & House, 1963; Recasens, 1991), the overall magnitude of carryover compared with anticipatory effects is a less reliable measure of the directionality of effects than is the relative degree of differentiation in midpoint value as a function of individual C^1 compared with C^2 contexts.

As shown in section 5.3, schwa displays a high degree of separability in F2 midpoint value as a function of both C^1 and C^2 place of articulation. With the exception of /ɪ/, the full vowels all show greater separability in midpoint value as a function of preceding consonantal place of articulation even though, in the case of /ɛ/ and /ʊ/, there is a stronger anticipatory effect in terms of the overall amount of perturbation to the vowel midpoint value.

5.4.1.3 F1

The total proportion of explained F1 variance for schwa is considerably lower than the total proportion of explained F2 variance. One possible interpretation for this difference in results, is that the lower context-dependency for F1 is reflective of an F1 target. That schwa should be specified for jaw position, albeit weakly, is to be expected given that some degree of jaw opening is inherent to its characterisation as a vowel. Keating (1990) describes an analogous situation for English vowels with respect to nasality. Because vowels in English are phonologically unspecified for nasality, they are characterised by wide windows for velum position and therefore accomodate most interpolations between consonants along this dimension. However, they will not permit a completely linear interpolation between two nasal consonants (characterised by maximal opening for velum position) since a slight raising of the velum is required in order to satisfy the [+oral] specification. However, beyond being characterised by an open constriction as opposed to a stricture of complete closure, there is no evidence in the present data to indicate a specific F1 target for schwa. As the coefficient of variance values demonstrate (see Table 5.2), schwa is actually more variable along F1 than along F2. With the exception of a small number of nasal contexts, the mean F1 trajectories for schwa also show interpolation through.

It is, therefore, more likely that the lower overall proportion of explained F1 compared with explained F2 variance reflects a higher proportion of random variability present for F1. It is also possible that temporal structure is more important for F1 than for F2 in which case formant values sampled at the durational vowel midpoint may be comparatively less representative of the F1 target. While the comparison of sampling points in Chapter 4 did not reveal any significant differences between F1 and F2 with respect to the relative performance of the midpoint, this possibility cannot be ruled out. Success in characterising vowels was judged in terms of the statistical separability of vowel distributions. For this, only absolute formant values obtained with each measurement method were used as input to the classifier. If, as Di Benedetto (1989) suggests, vowels vary inherently with respect to the time at which F1 maxima are attained and if, as the results of the present survey indicated, temporal structure also varies as a function of context, in order to do full justice to the formant maxima method in a comparative evaluation of its performance (either across vowels or across formant frequencies), additional temporal and contextual information would also need to be supplied. The fact that several vowels display higher levels of variability for F1 than for F2 and that the R^2 values for F1 are lower than the corresponding R^2 values for F2 in all cases are consistent with a general explanation along these lines.

Despite the difference in the overall R^2 value obtained for F1 and F2, the 74% explained F1 variance nevertheless represents a significantly high level of context dependency. In view of this and, in view also of the high F1 variability for schwa and the linearity of the schwa mean F1 trajectories, it may be concluded that schwa is also highly unspecified along this formant dimension. As in the case of F2, the full range in F1 value for schwa does not cover the maximum frequency range possible for F1 across all vowels. Specifically, there are comparatively few tokens characterised by very high F1 values. However, as before, the apparent constraint on variability in schwa midpoint value is less reflective of an inherent target specification than of contextual limitations.

There is also generally less interaction of effects for F1 than for F2. However, as shown in Figure 5.36, schwa displays a higher correlation between F1 onset and offset value than any other vowel. Results of the analyses of variance and Tukey multiple comparison of means tests also show a relatively higher degree of differentiation in schwa midpoint value as a function of both C^1 and C^2 place or manner effects than the other vowels. Thus, although less conclusive than in the case of F2, the range and directionality of coarticulatory effects for F1 are not inconsistent with the targetless hypothesis.

5.4.1.3 Comparison of results with other studies

The amount of explained variance in schwa midpoint values for both F1 and F2 is considerably higher in the present study than that reported by Van Bergem (1994) for schwa in ˈVCəC and CəˈCV nonsense sequences. For F1, Van Bergem reports 29% and 38% and for F2, 72% and 79% explained variance for ˈVCəC and CəˈCV sequences respectively. However, the present result for F2 is comparable with the 87-89% explained F2 variance (obtained for three speakers) reported by Kondo (1995) for schwas in VCəCV sequences excised from meaningful, connected speech data. The lower percentage of explained variance in Van Bergem's study may be attributable, therefore, to the fact that nonsense words and words in carrier phrases tend to be pronounced more carefully than more natural speech data (Koopmans van Beinum, 1980;

Krull, 1989; Lindblom & Moon, 1988). More careful pronunciation typically results in longer durations and less coarticulation between segments (Engstrand, 1988). While this is a plausible explanation for the observed difference in results, allowing for style dependent variation in the amount of context-dependency shown by schwa is potentially problematic for the targetless analysis. If schwa is completely unspecified along F1 and F2, it should show similar levels of context-dependency across different speaking rates and styles.

In an attempt to resolve this apparent anomaly, it is proposed that schwa is targetless insofar as it represents an empty time-slot in a manner similar to that proposed by Anderson (1982) for French schwa, but that different phonetic consquences are manifest during this time-slot depending on variation in factors such as speech style and rate and therefore also in duration and amount of coarticulation. (In contrast to Anderson's phonological analysis, the phonetic structure is assumed here to be implemented at surface level.) In circumstances of short segmental durations and/or a relatively high degree of coarticulation between segments, the articulatory and acoustic interval corresponding to schwa will reflect interpolation through. In more careful or artificial styles of speech, an increase in the duration of the schwa interval and/or a general decrease in the amount of coarticulation between adjacent specified segments may result in articulatory relaxation or drift during schwa giving rise to formant values characteristic of a general relaxation position for the tongue.

This analysis is consistent with predictions made by Browman & Goldstein's (1986; 1989; 1992) model of articulatory phonology. According to Browman & Goldstein (1992), passive changes in the value of tract variables such as degree and location of tongue body constriction during intervals when they are not under active gestural control, occur through coupling with a variable that is under active control and/or an articulator-specific 'neutral' or 'rest' regime. Thus, with respect to their own data (see Chapter 2, section 2.3.2.3 for full description), Browman & Goldstein (1992) consider the possibility that the warping in the tongue trajectory they observe for medial schwa in

/pipəˈpipə/ sequences, represents movement toward "an overall average or neutral tongue position" (p.42). However, on the basis of the results of tests involving computer simulated data, they reject this analysis and argue instead in favour of an active target which is specific to schwa.

In the final simulation they perform, in which no tongue body gesture is active during the schwa interval, Browman & Goldstein report a lowering of the tongue trajectory for all utterances including the /papəˈpapə/ sequence. In the X-ray data, this sequence was characterised by a raising of the tongue-body during schwa. They attribute the general lowering of the tongue body for the simulated data to the fact that

> ...the neutral position contributing to the tongue-body movement was that of the tongue-body articulators rather than that of the tongue-body tract variables; consequently the dip was relative to the jaw, which, in turn, was lowering as part of the labial release. (p. 54)

On the basis that the neutral positions for individual articulators produced a different articulatory effect from the neutral position "defined in the space of tract variables" (i.e. location and degree of tongue body constriction), Browman & Goldstein conclude that "a target position for schwa was specified, although this target is completely predictable from the rest of the system; it corresponds to the mean tongue tract-variable position for all the full vowels" (p.56).

In his comments on their paper, Barry (1992, same volume) challenges Browman & Goldstein's conclusion. He claims that, within the articulatory phonology framework, it is not possible to distinguish between a hypothesised targetless schwa defined as an empty time slot from the schwa target they propose. Firstly, because "a phonologically targetless schwa could still not escape the residual dynamic forces of the articulatory muscular system, i.e. it would be subject to the relaxation forces of that system" and secondly, in view of the 'coordinative' assumption of articulatory control implicit in a task-dynamic system "which argues against one gesture being relaxed independent of other relevant gestural vowel parameters" (p.66). Given that "any muscle in a functional system can only be accorded a "neutral" or "relaxation" value as a function of the forces brought to bear on it by other muscles within the system", Barry concludes that

> the mean vowel specifying the schwa "target" used by Browman and Goldstein is identical with the relaxation position of the vocalic functional system since it reflects the balance of forces between the muscle tension targets specified for all the other vowels within the system. (p.66)

Thus, accepting the concept of a general relaxation position and, barring Browman & Goldstein's distinction between this and a neutral position specific to schwa, the targetless analysis is workable. Characterisation of schwa as a inherently unspecified time-slot with different phonetic reflexes depending on variation in the timing and inter-articulator alignment of gestures has a number of advantages. In addition to accomodating potential differences in absolute degree of context dependency observed for schwa across different speech samples, this analysis also allows for potential differences in the articulatory and acoustic patterns for schwas occupying different temporal positions within an utterance. For example, utterance-final schwas are likely to exhibit different coarticulatory effects from schwas surrounded by context segments. Presumably, interpolation through can only occur in the case of the medial schwas. Utterance-final schwas may be expected to show carryover effects from the preceding segment(s) and relaxation towards the rest position in anticipation of the following silence. This is consistent with Nord's (1974) observations for final compared with non-final unstressed vowels in Swedish: "unstressed vowels coarticulate strongly with context in non-final position ... and in final position with a rest-like position corresponding to an acoustic schwa vowel" (pp.153-154). Furthermore, as Barry (1992) points out, the idea of a general relaxation position also accords well with the differences in neutral vowel quality observed across different languages. Presumably the same also applies in the case of dialects with different vowel systems. For example, the neutral vowel in many Scottish dialects is characterised by a raised quality relative to schwa in SBS.

One situation in which it is difficult not to envisage an active gesture or 'target' for schwa is when it is pronounced in isolation. However, in this case, it can be argued that the 'target' is the same as the articulatory configuration and acoustic pattern characteristic of the general relaxation position, but that its attainment is 'intentional' as opposed to passive. It is also possible that highly artificial speech data of the kind used by Browman & Goldstein (1992) results in a similarly 'targeted' schwa in view of the relatively high degree of conscious introspection the task of producing an unnatural sequence in experimental conditions is likely to involve.

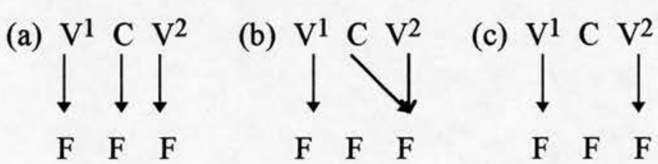
One question that emerges following Barry's critique of their paper, is how to explain the different articulatory patterns Browman & Goldstein observe for the /papəˈpapə/ sequence in the computer simulated and X-ray data. It is possible that these reflect methodological problems inherent in the modeling of the simulated data. Browman & Goldstein note that it was not possible to model the differential patterns of motion displayed by the middle compared to the rear of the tongue body in the X-ray data. While they report a general similarity between the two sets of data, some differences were apparent such as generally more extreme positions and lower overall variability for the simulated data. However, it is also possible that, although mistaken in distinguishing between a general relaxation position and neutral position specific to schwa, Browman & Goldstein are correct in their analysis of an active (as opposed to passive) target for /ə/ in *their* data. Since /ə/ produced in isolation or in artificial nonsense syllables is arguably not the same phonological entity as /ə/ which occurs in meaningful connected speech, the question of whether or not Browman & Goldstein's data reflect an active or passive schwa gesture is largely academic. For schwa tokens produced in the current connected speech data-base, there is no evidence of a specific F1 or F2 target.

5.4.1.4 Relative influence of adjacent vowels and consonants

The results of the regression analysis for both F1 and F2 show that, with respect to degree of coarticulatory influence, the temporal position of context segments in relation to schwa is more important than either their vocalic or consonantal status.

The absence of significant vowel-to-vowel effects in the present data, except in those cases where schwa is immediately adjacent to another vowel (i.e. in the CVəCC and CVəCV sets) and likewise of consonant-to-vowel effects across an intervening vowel, may be attributed to the unrestricted nature of the speech material and to extensive consonant-vowel interactions. Keating (1988) describes three possibilities with respect to the type of vowel-consonant interactions likely to occur in a V^1CV^2 sequence. These are illustrated schematically in Figure 5.47 below. If as in (a) V^1, C and V^2 each have their own specification for a given feature (F), vowel-to-vowel effects in either direction will be blocked by the consonant. If as in (b) C has assimilated the feature value of V^2, the vowel-to-vowel effect from V^2 to V^1 will in fact be a vowel-to-consonant-to-vowel effect. There will be no effect of V^1 on V^2 because C's acquired feature value will serve to block interaction. If as in (c) C has no specification of its own and remains unspecified, vowel-to-vowel effects will occur freely in both directions.

Figure 5.47: Consonant-vowel interactions in a VCV sequence.



The small effect of a next adjacent vowel or consonant observed for some of the data sets may reflect a weakening rather than a complete blocking of effects. Alternatively, it may also represent instances of (c).

In addition to the potential blocking effect of intervening consonants, the small vowel-to-vowel effect in the present data may also reflect the high variability in formant frequency value for the context vowels which include both stressed and unstressed tokens, the diphthongs and a high proportion of /ɪ/ and /ə/ tokens. Other studies which report strong vowel-to-vowel effects (Magen, 1989; Fowler, 1981; Browman & Goldstein, 1992) use a restricted set of context vowels which, owing to the limited consonantal context and the controlled nature of the speech material (isolated nonsense syllables), are unlikely to have shown much variation from their target values. They also minimise consonant-vowel interactions by examining vowel-to-vowel effects across an intervening bilabilal stop.

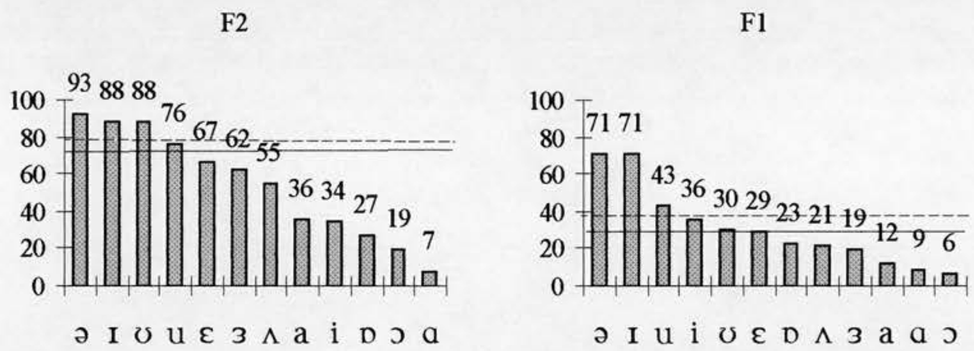
Van Bergem (1994) examines schwa in a wider range of consonantal contexts. In general, his results also indicate that temporal position in relation to schwa is more important than either the vocalic or consonantal status of the context segments. Although a greater vowel than C^1 effect is noted for schwa in $C^1əC^2V$ words (24% explained F2 variance compared with 21% for C^1 and 26% for C^2), the C^1 and C^2 effects are both larger than the vowel effect for schwa in $VC^1ə\ C^2$ words (15% and 48% compared with 9% respectively). The large difference in relative C^1 and C^2 contribution in the $VC^1əC^2$ words suggests generally greater anticipatory than carryover coarticulation which may also account for the relatively strong vowel influence in the $C^1əC^2V$ words. The stronger vowel-to-vowel effect reported by Van Bergem might also be attributable to the fact that the context vowels in his data were restricted to the point vowels /i, a:, u/ and that these were also stressed. In representing the extremes of vowel quality, the point vowels are likely to exert the strongest coarticulatory influence of all the vowels. As Fowler (1981) demonstrates, stressed vowels also exert a stronger coarticulatory influence than unstressed vowels.

5.4.2 Inherent variability: full vowels

The proportion of variance explained by context alone (i.e. excluding the unique contribution made by duration) is shown graphically for all the vowels in Figure 5.48.

For comparison, the proportion of explained variance for schwa reported by Van Bergem (1994) is also shown, indicated by the two horizontal lines. The dotted line represents the results for schwa in ˈVCəC words and the unbroken line, the results for schwa in CəˈCV words.

Figure 5.48: Proportion of variance explained by preceding and following context.



As Figure 5.48 demonstrates, the full vowels /ɪ, ɛ, ʌ, u, ʊ/ and /ɜ/ all display high levels of context dependency along F2. With the exception of /i/, all vowels display lower levels of context-dependency along F1. The greater or comparable proportion of explained F2 variance for /ɪ, ɛ, u, ʊ/ and of explained F1 variance for /ɪ, u, i, ɛ/ to the levels reported for schwa by Van Bergem is indicative of the difference in amount of coarticulation that can occur across speech samples as a function of factors such as variation in speech rate and style. It thereby serves to highlight the importance of assessing schwa's context-dependency within a comparative framework.

5.4.2.1 Front versus back vowels

Recasens (1991) predicts a greater magnitude of anticipatory coarticulation and hence a greater range of effects for schwa and for back vowels than for front vowels on the grounds that the tongue body is subject to a lesser degree of constraint during their production (see section 2.3.2.1 for a full discussion). While this prediction is supported in the present data with regard to schwa compared with the full vowels, the results do not support his predictions for back vowels compared with front vowels. Recasens reports higher levels of variability for the back vowels /u/, /o/, /ɔ/ and /a/ in Catalan than for the front vowels /i/, /e/ and /ɛ/. The present results accord with Recasens' data in demonstrating relatively low F2 context-dependency for the tense, front vowel /i/. However, they also show relatively high context-dependency for the lax, front vowels /ɪ/ and /ɛ/ in comparison with the back vowels /ɑ, ɒ, ɔ/ and /ʌ/. (As in Recasens' data, relatively high levels of context-dependency are also observed for the back vowel /u/. However, in the present study, /u/ is arguably disqualified from the front-back comparison owing to its fronted realisation.)

In general, the back vowels show lower levels of anticipatory coarticulation and consequently also a less extensive range of coarticulatory effects than the front vowels. This result is attributable, in part, to differences between the two vowel classes with respect to which contexts exert the most coarticulatory influence and to an unequal distribution of these contexts in the present data. For example, in the case of /i/, /ɛ/ and /a/ the observed anticipatory effect is almost entirely due to the lowering influence of the velarised lateral /ɫ/. The effects from velarised /ɫ/ on back vowels are small in comparison. Recasens explains the difference in effect for the two classes of vowel in terms of gestural compatibility. The coarticulatory effect of velarised /ɫ/ is particularly strong for front vowels because it involves a low back tongue position which conflicts with the raised and fronted tongue position characteristic of front vowels. It has a small effect on back vowels because they are similarly characterised by a low back tongue position. In Recasens' data, larger effects on back vowels were produced by palatal and dento-alveolar consonants. The high tongue position required for these consonants caused an increase in the degree of linguapalatal contact and a corresponding increase in F2 frequency for the vowel.

In the present data, the lack of a significant effect from a following palatal may be attributed to the small number of following palatal contexts (between five and ten tokens for each back vowel). However, the lack of a significant effect from a following alveolar or dental consonant cannot be attributed to small sample size. Nor can it be attributed to an interaction of effects such that the raising influence of a following alveolar/dental is cancelled out by the lowering influence of a preceding low locus consonant. The results of the two-way analyses showed no significant C^1 x C^2 interaction for any of the back vowels. A higher degree of differentiation in offset than in midpoint value as a function of C^2 place of articulation (see section 5.3.1) confirms that the relatively low level of anticipatory coarticulation displayed by /ɒ/ and /ʌ/ and the lack of an observed anticipatory effect for /ɑ/ and /ɔ/ are not solely a question of contextual limitations but also reflect properties of the vowels.

Thus, allowing for differences in speech material and methodology and also for differences in coarticulatory strategies between speakers (Nolan, 1985), the present data do not appear to support Recasens' (1991) proposals. In view of the relatively low variability displayed by the front vowels /i/ and /a/ and the back vowels /ɑ/, /ɔ/ and /ɒ/, and, conversely, the relatively high context-dependency displayed by the lax, back vowel /ʌ/ and the lax, front vowels /ɪ/ and /ɛ/, in the current data, it is not possible to distinguish vowels in terms of inherent variability on the basis of their front/back classification. The apparent discrepancy between the present results and Recasens' findings is further discussed in section 7.2.4.

5.4.2.2 Tense versus lax vowels

A strict division between tense and lax vowel sets with respect to inherent variability is also not possible owing to the relatively high degree of context-dependency displayed by the tense vowel /u/ and the relatively low degree of context-dependency displayed by the lax vowels /a/ and /ɒ/. However, it is the case that within each front/back category, the tense vowels and the long, lax vowel /a/ are more robust than the short, lax vowels. Thus, lower R^2 values are obtained for /i/ and /a/ compared with /ɪ/ and /ɛ/ and for /ɑ/ and /ɔ/ compared with /ɒ/ and /ʌ/. This is true for both F2 and F1. /u/ and /ɜ/ also display lower overall context-dependency along F2 than /ʊ/ and /ə/. However, while /ɜ/ is also more robust than /ə/ along F1, /u/ displays a higher R^2 value for F1 than /ʊ/.

In general, the greater context-sensitivity of lax compared with tense vowels is reflected in the level of F2 variability displayed by each vowel although overall degree of context-dependency and overall range in midpoint value are not in one-to-one correspondance. Again, a strict division between tense and lax vowel sets is not possible, this time owing to the relatively high degree of variability displayed by /u/ and /ɔ/. Both vowels show a greater range in midpoint value than either of the lax vowels /ɛ/ and /ʌ/. However, in each case, the tense vowels are less variable than their corresponding lax counterparts /ʊ/ or /ɒ/. With respect to F1, /i/ and /u/ show the highest variability after schwa and /ɪ/. Otherwise the lax vowels show greater variability than the tense vowels.

The results for the mid-high, lax vowels /ɪ/ and /ʊ/ indicate that, as in the case of schwa, these vowels are also unspecified for tongue position. Both vowels show a comparable range in F2 midpoint value to that displayed by schwa and similarly little difference in the range of midpoint compared with the range in onset or offset value. This is illustrated graphically by the distribution and linearity of the mean formant trajectories for each vowel. The high overall context-dependency shown by /ɪ/ and /ʊ/ is confirmed in the regression analyses. Accepting the results of the analysis in which the extreme outliers for /ɪ/ are excluded (see section 5.2.3.1), both vowels show a comparable level of context-dependency to schwa.

As the results of the analysis of variance and post hoc multiple comparison of means tests demonstrate, /ɪ/ also patterns with schwa with respect to the directionality of effects, both vowels showing a greater parity in the amount of carryover compared with anticipatory coarticulation than the other vowels. /ʊ/, however, patterns with the full vowels showing a greater degree of differentiation in midpoint value as a function of preceding consonant place of articulation. /ʊ/ also shows a relatively low level of overall variability and context-dependency along F1 while /ɪ/ displays a comparable degree of F1 context-dependency to schwa.

The relatively low contextual variability along F1 for /ʊ/ may reflect its inherent specification for lip-rounding. The increase in overall tract length caused by lip-rounding serves to lower both F1 and F2. However, this lowering effect is arguably more likely to be cancelled out in the case of F2 through variation arising as a function of variation in tongue position. /ɯ/ also displays relatively low context-dependency for F1 compared with F2 although the difference in the R^2 value for F1 and F2 is less than in the case of /ʊ/. It is also possible that the lower F1 variability for /ʊ/ reflects a more limited range of contextual influence.

Less systematic effects for F2 may also be attributed to the influence of lip-rounding. Because lip-rounding affects the length of the front oral cavity, in connected speech it is likely to confound the raising and lowering effect of adjacent consonants on the vowel second formant values and thereby cancel out differences between individual place categories. Both vowels show relatively little differentiation in onset value as a function of preceding consonantal place of articulation. They also display greater overall variability in F2 midpoint value than either schwa or /ɪ/.

5.5 Duration

A significant, unique contribution of duration to the F2 prediction for all vowels excepting /ɯ/ and /ɜ/, indicates a process of duration-dependent centralisation operating in addition to the process of contextual assimilation. However, in most cases, the amount of variance in F2 midpoint value uniquely accounted for by duration is negligible in comparison with the proportion accounted for by context (see section 5.2.3.3). The only vowels which show a relatively large independent durational effect are the high, front vowel /i/, the mid-low, back vowels /ɒ/ and /ɔ/ and the low, back vowel /ɑ/. In absolute terms, the proportion of variance in /ɑ/ F2 midpoint value uniquely accounted for by duration is considerably lower than the proportion accounted for in the case of /i, ɒ, ɔ/. However, given the lower overall variability in /ɑ/ F2 midpoint value, the relative contribution of duration is comparable. Predictably, in the case of each vowel, shorter durations generally imply more central values.

Evidence of an independent process of centralisation conflicts with Lindblom's (1963) Undershoot Hypothesis (see Introduction to Chapter 6). It is, however, consistent with a gestural overlap account in which vowels are characterised as ranges or regions rather than as points in articulatory and acoustic space. Given that /i/, /ɑ/, /ɒ/ and /ɔ/, which represent the extremes in front/back tongue position (and hence the upper and lower limits in F2 value), are generally less compatible with context than vowels characterised by less extreme articulatory configurations (and hence more central F2 values), these vowels arguably have a wider 'target' area over which to vary without being subject to overlap. At the acoustic level, this results in values which although more central than the most extreme values possible, are still higher or lower than the majority of consonantal locus values. For example, in the case of /i/, tokens which display F2 midpoint values of 2000 Hz are centralised relative to tokens which display midpoint values in the region of 2200 Hz. However, the only consonant likely to be characterised by a similarly high value is the palatal glide /j/.

With respect to F1, the larger independent durational effect observed for the the low and mid-low vowels (excepting /ɔ/) relative to the high and mid-high vowels may similarly be attributed to the fact that these vowels have a greater distance over which to vary without being subject to overlap. The lack of a significant durational effect for the mid-low, back vowel /ɔ/ may reflect its generally low variability in F1 midpoint value. The role of duration in conditioning vowel quality is further discussed in Chapters 6 and 7.

5.6 Summary

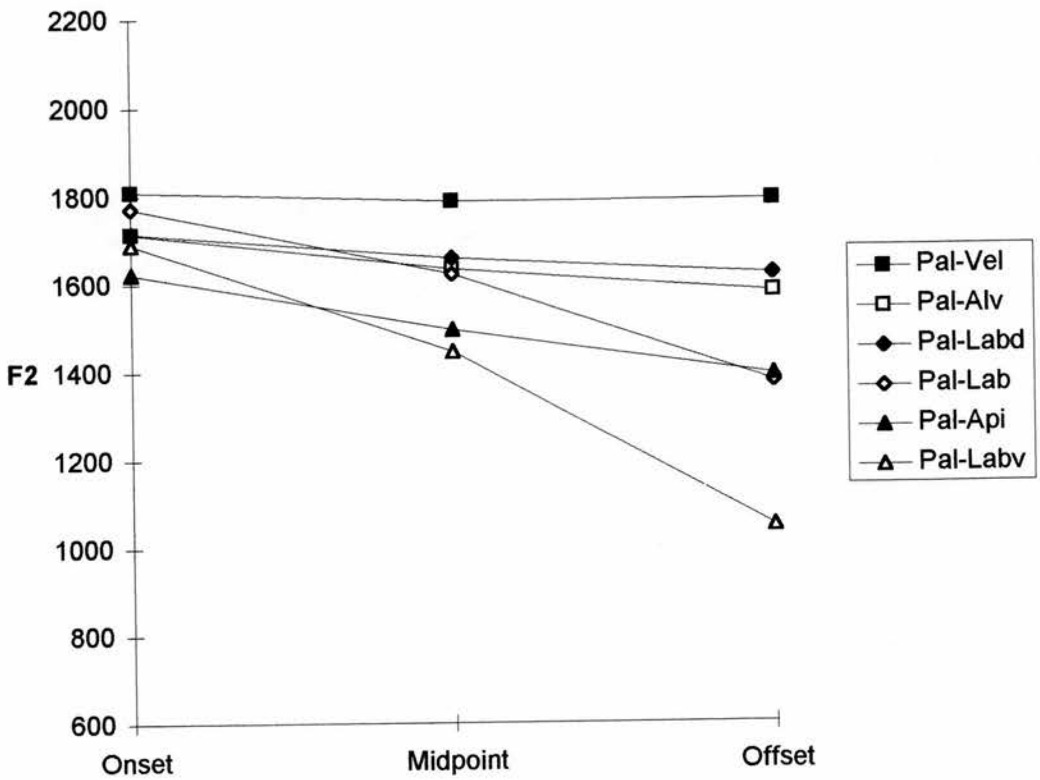
In this chapter, vowels are compared with respect to the overall variability and patterns of variability they exhibit as a function of context. As predicted, schwa exhibits a higher overall degree of context-dependency than any of the full vowels. The high proportion of F1 and F2 variance explained by context, the interaction of context effects through schwa and the high degree of differentiation in schwa midpoint values as a function of both preceding and following consonantal place of articulation strongly support the theory that /ə/ is unspecified for tongue and jaw position.

High variability and high levels of context-dependency are also displayed by /ɪ/, /ʊ/ and /u/. In the case of /ɪ/, the systematicity and interaction of context effects are such to suggest that it is also highly unspecified along F1 and F2. While results for /ʊ/ and, to a lesser extent, /u/, also indicate a high degree of underspecification for F2, there is evidence of a target for F1. In the case of both vowels, effects are less systematic than those observed for schwa and /ɪ/. This may be attributable to variability in degree of lip rounding and to the interaction of effects due to variation in both tongue and lip position.

In contrast to observations in the literature for Catalan (Recasens, 1991) and American English (Syrdal & Gopal, 1986), the present data do not support the view that back vowels are generally less robust than front vowels. Rather, the results indicate a hierarchy of inherent robustness in which the peripheral vowels /i, a, ɑ, ɔ/ and /ɒ/ show greater resistance to context effects than the more central vowels /ɪ, ɛ, ɜ, ʌ, ʊ/ and, in the present data, /u/. These results and their implications for an account of vowel reduction are discussed more fully in Chapter 7.

Figure 5.1: Mean second formant trajectories for schwa - preceding/following context = palatal

5.1a: preceding context



5.1b: following context

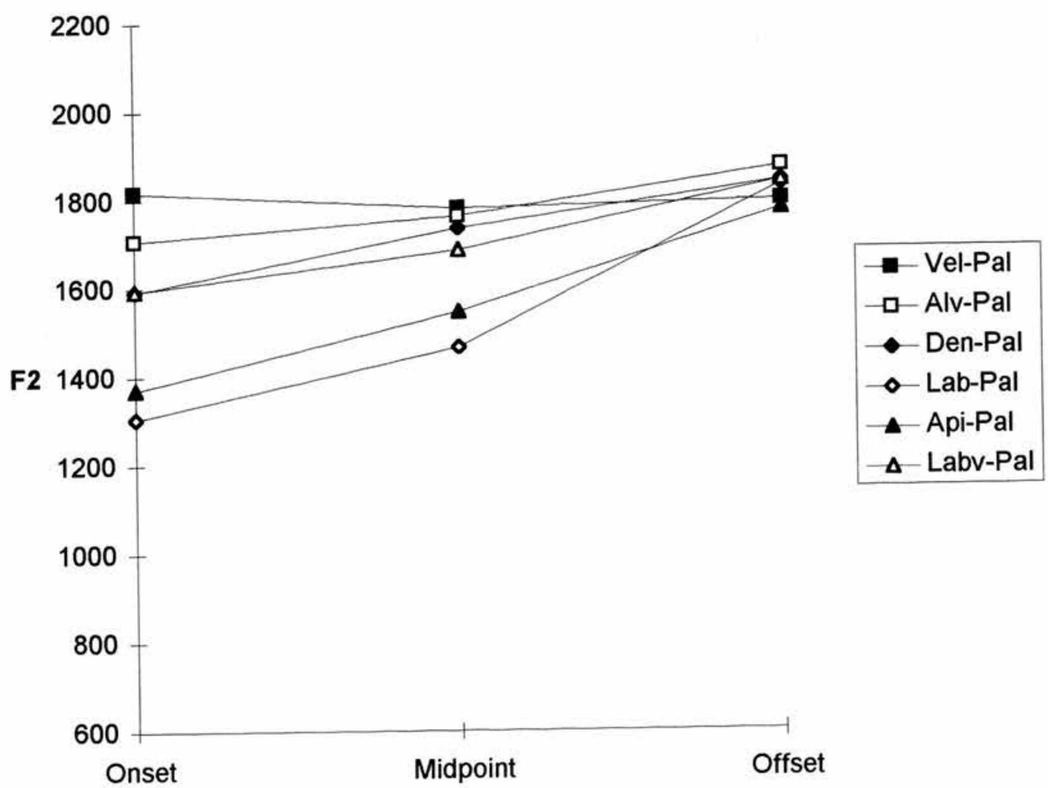
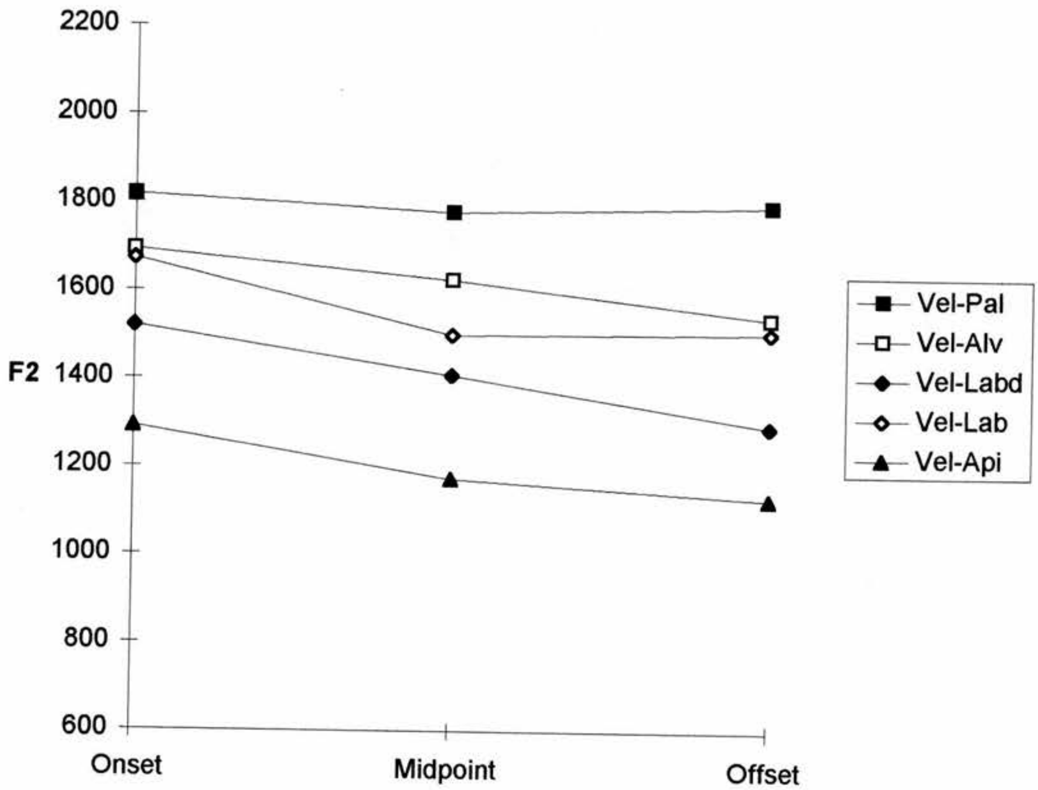


Figure 5.2: Mean second formant trajectories for schwa - preceding/following context = velar

5.2a: preceding context



5.2b: following context

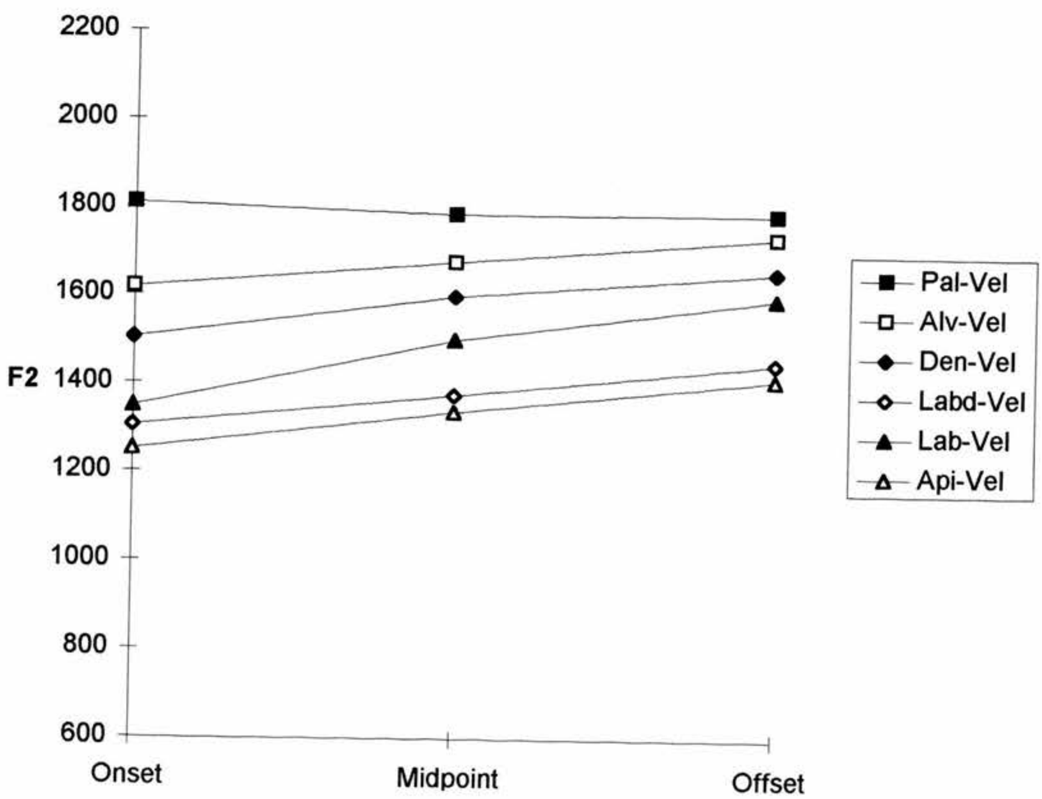
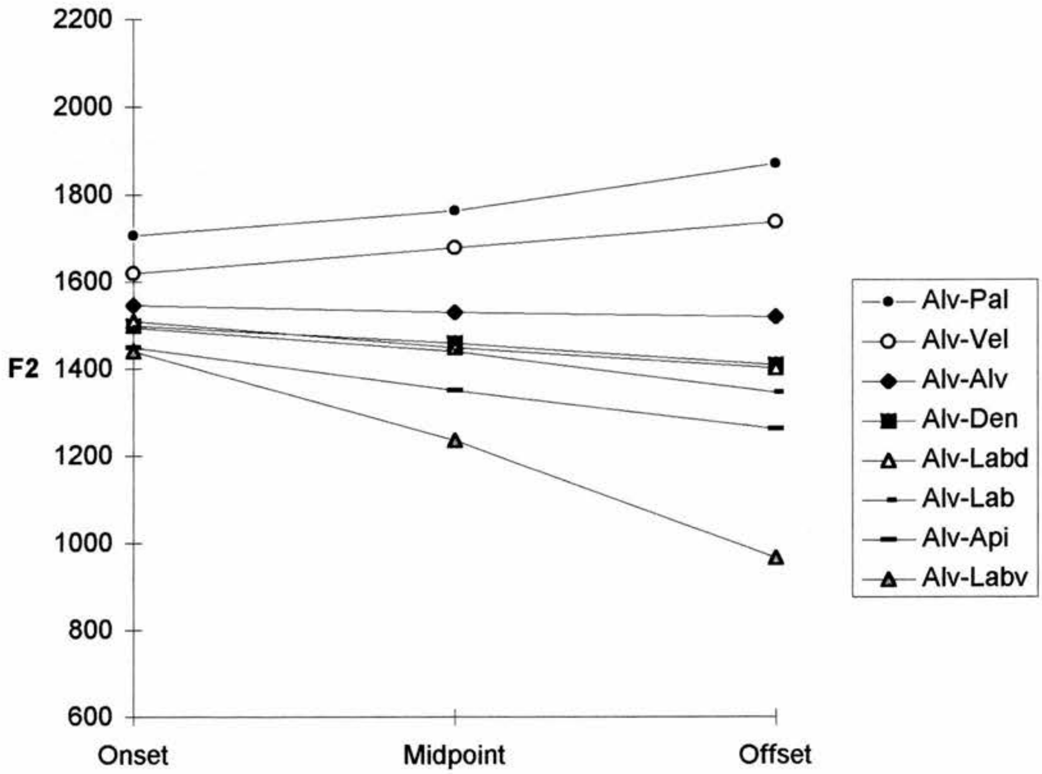


Figure 5.3: Mean second formant trajectories for schwa as a function of preceding consonant place of articulation

5.3a: alveolar



5.3b: dental

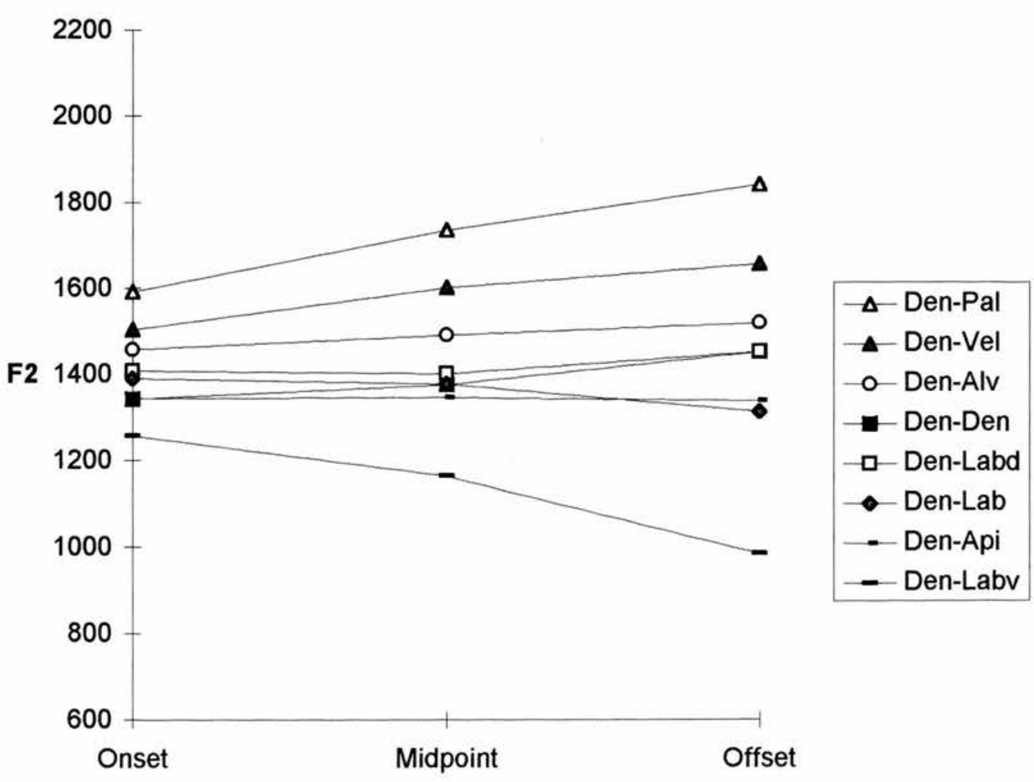
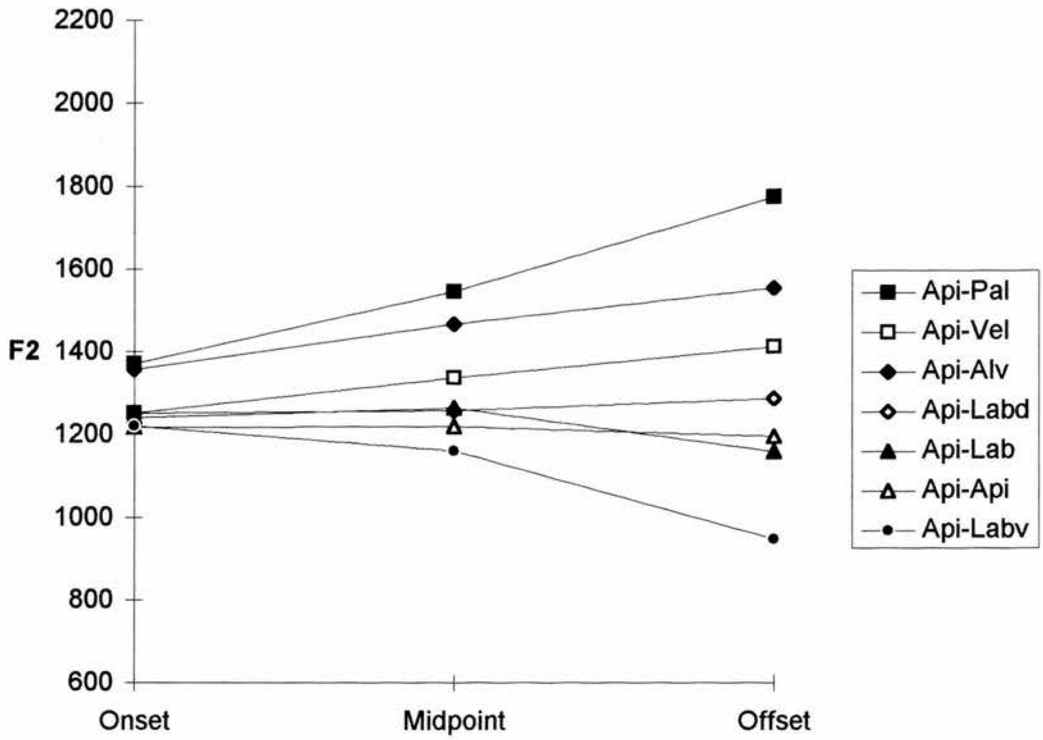


Figure 5.3: Mean second formant trajectories for schwa as a function of preceding consonant place of articulation (continued)

5.3e: apical



5.3f: labio-velar

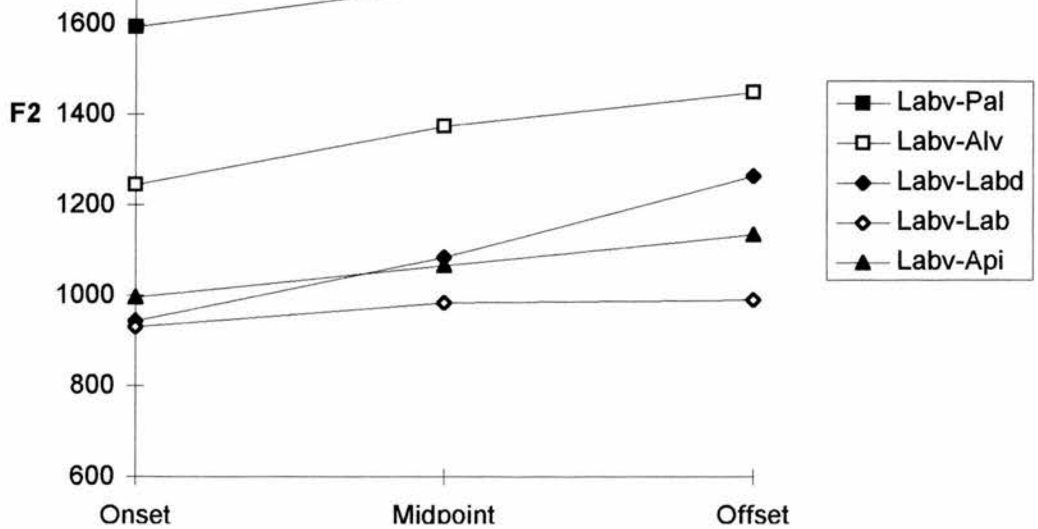
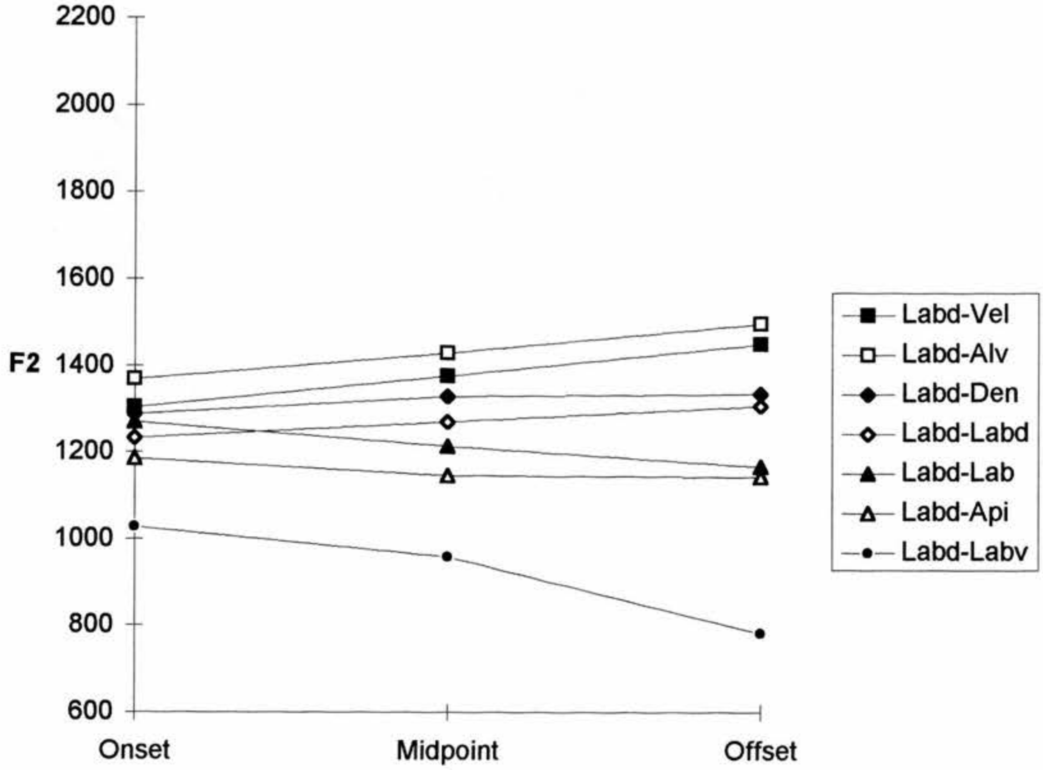


Figure 5.3: Mean second formant trajectories for schwa as a function of preceding consonant place of articulation (continued)

5.3c: labio-dental



5.3d: labial

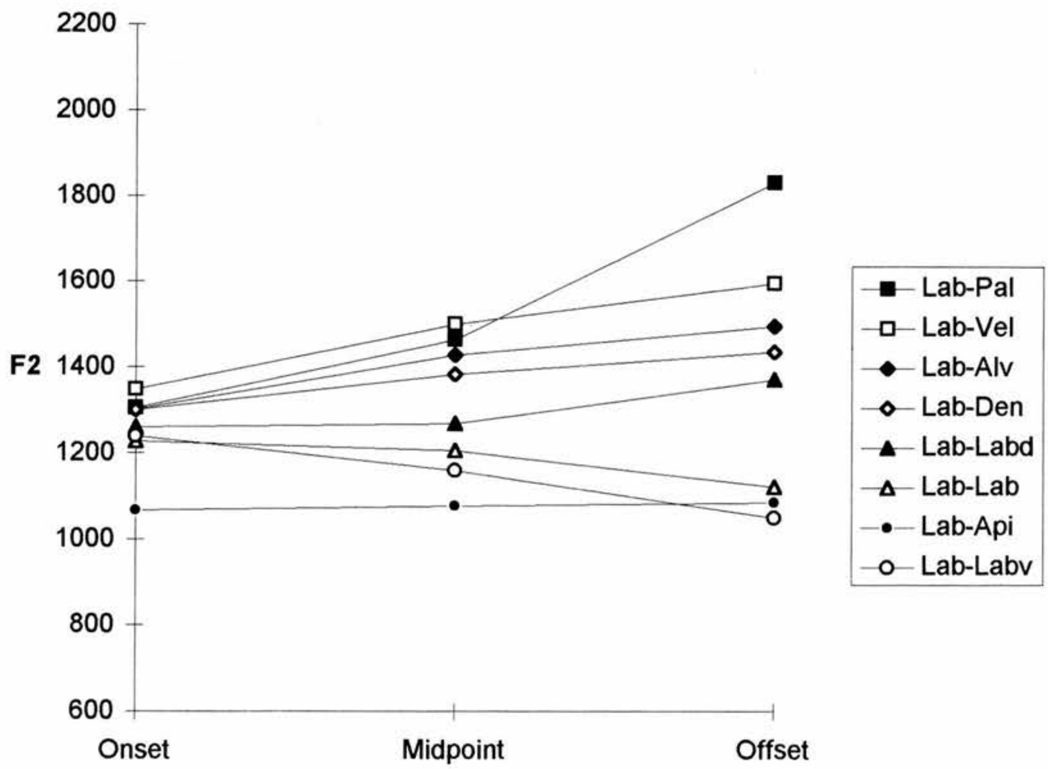
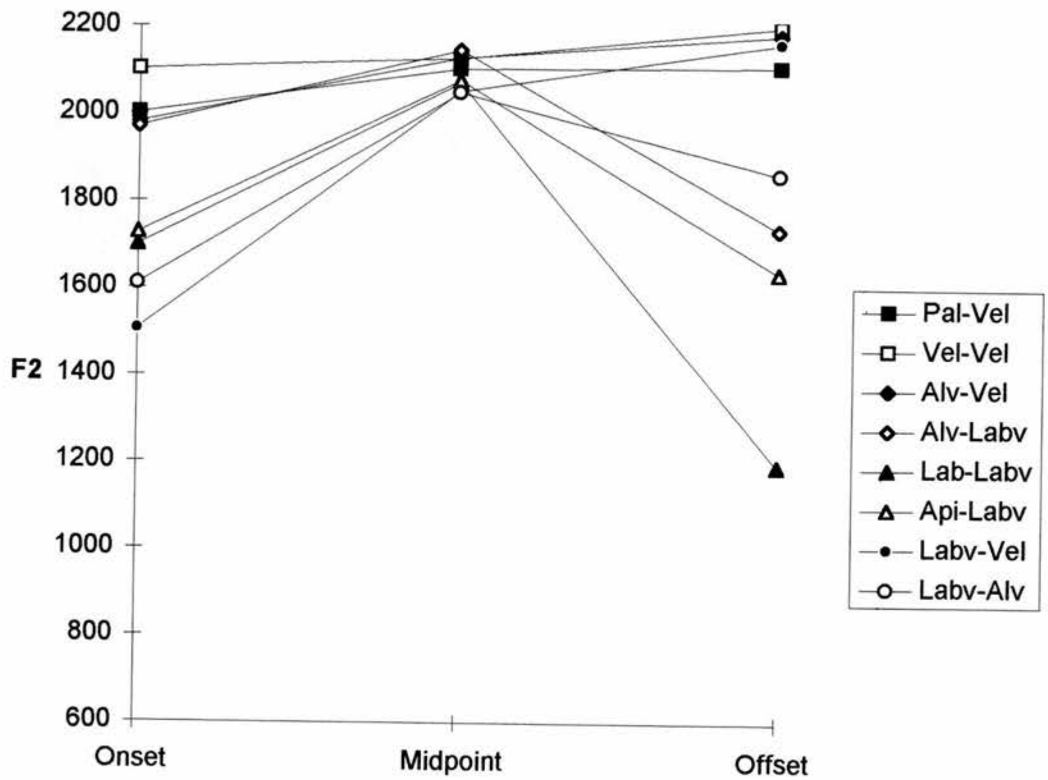


Figure 5.4: Mean second formant trajectories for /i/

5.4a: high deviation values



5.4b: low deviation values

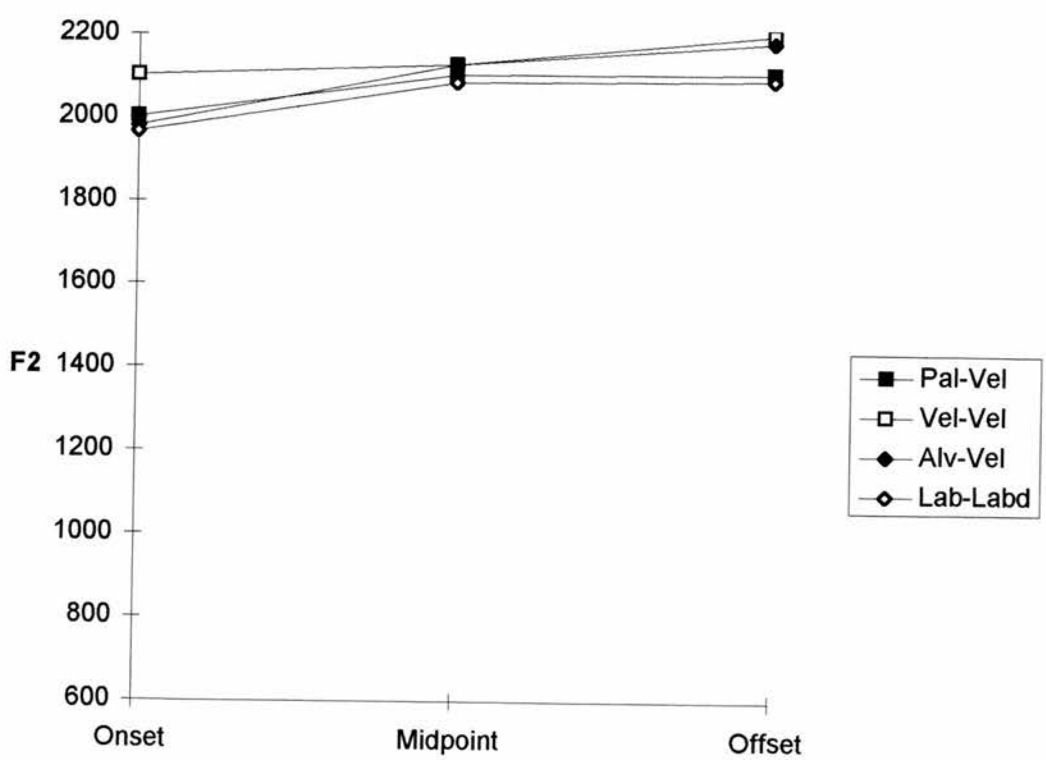
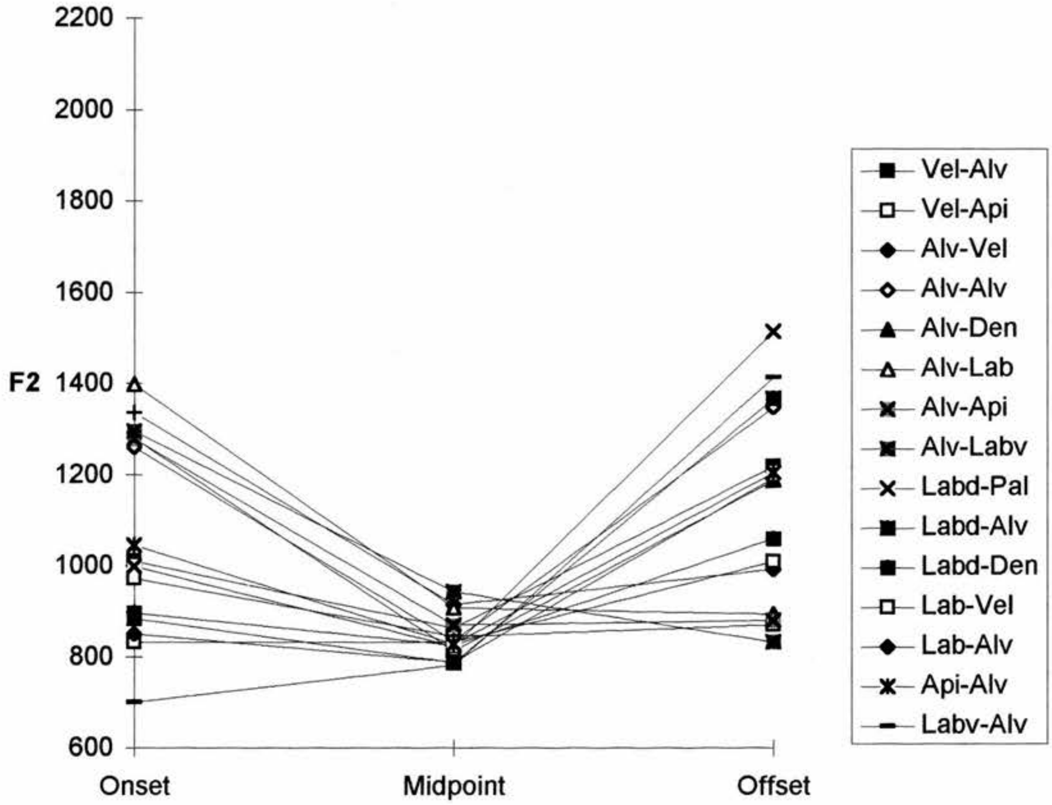


Figure 5.5: Mean second formant trajectories for /ɔ/

5.5a: high deviation values



5.5b: low deviation values

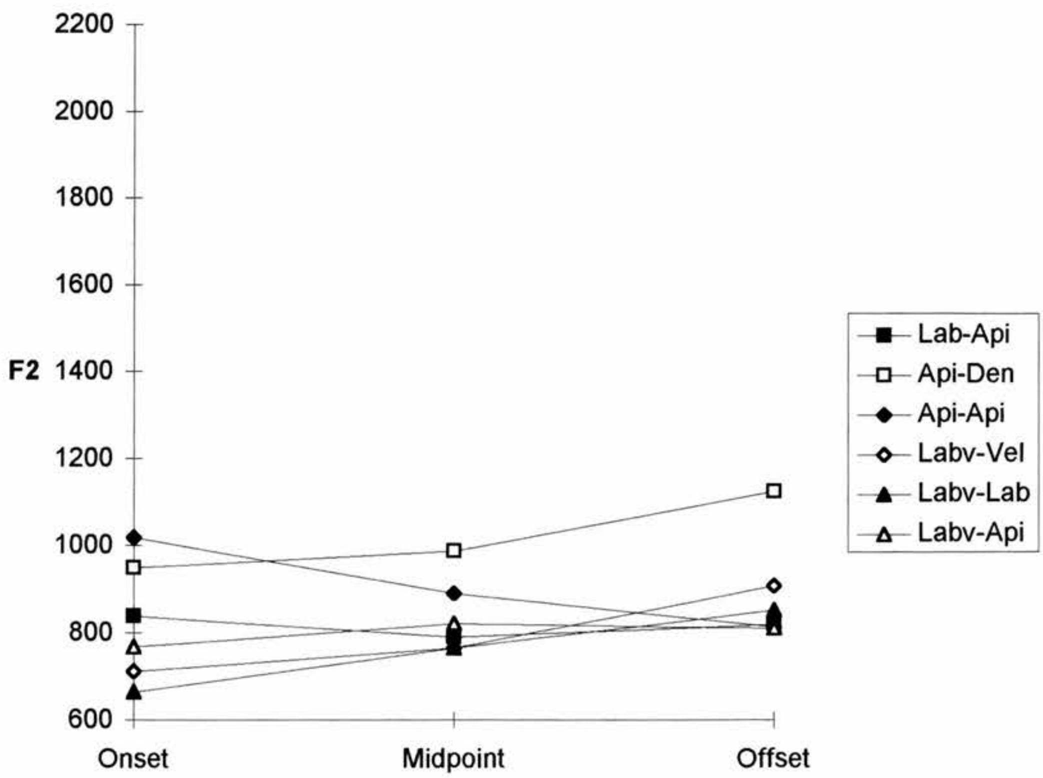


Figure 5.6: Mean second formant trajectories for schwa - symmetrical contexts

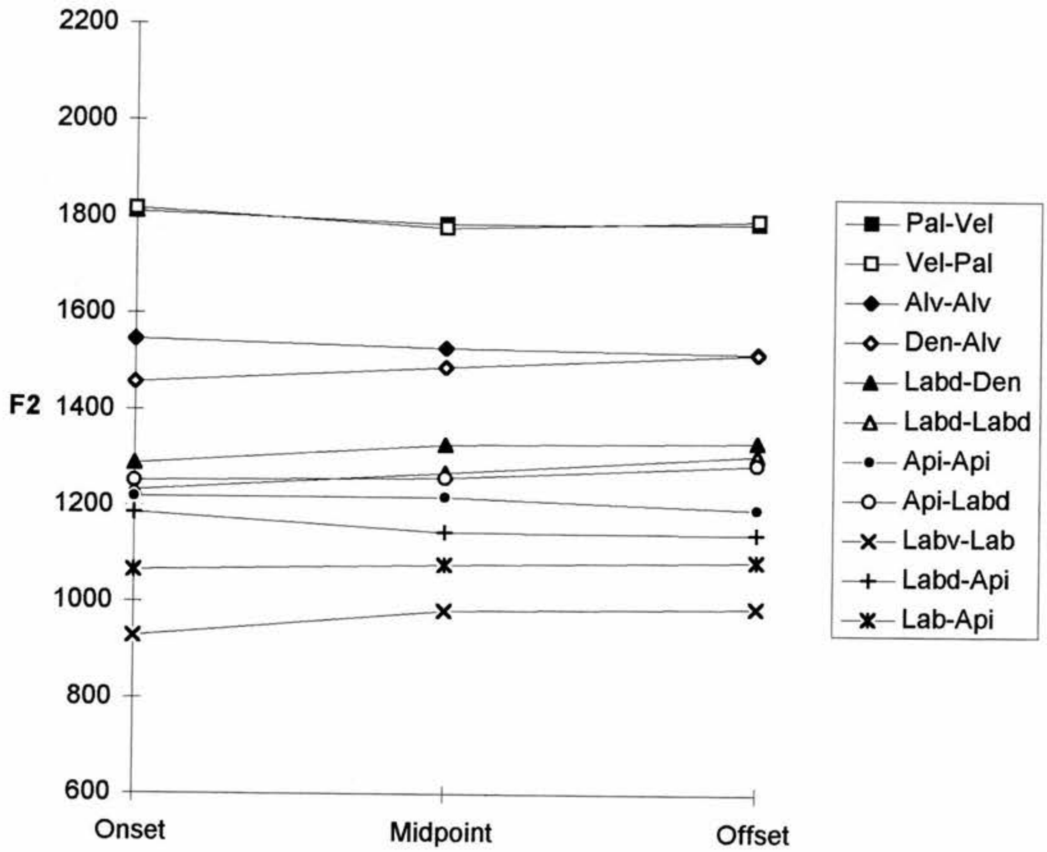


Figure 5.7: Mean second formant trajectories for schwa - high deviation values

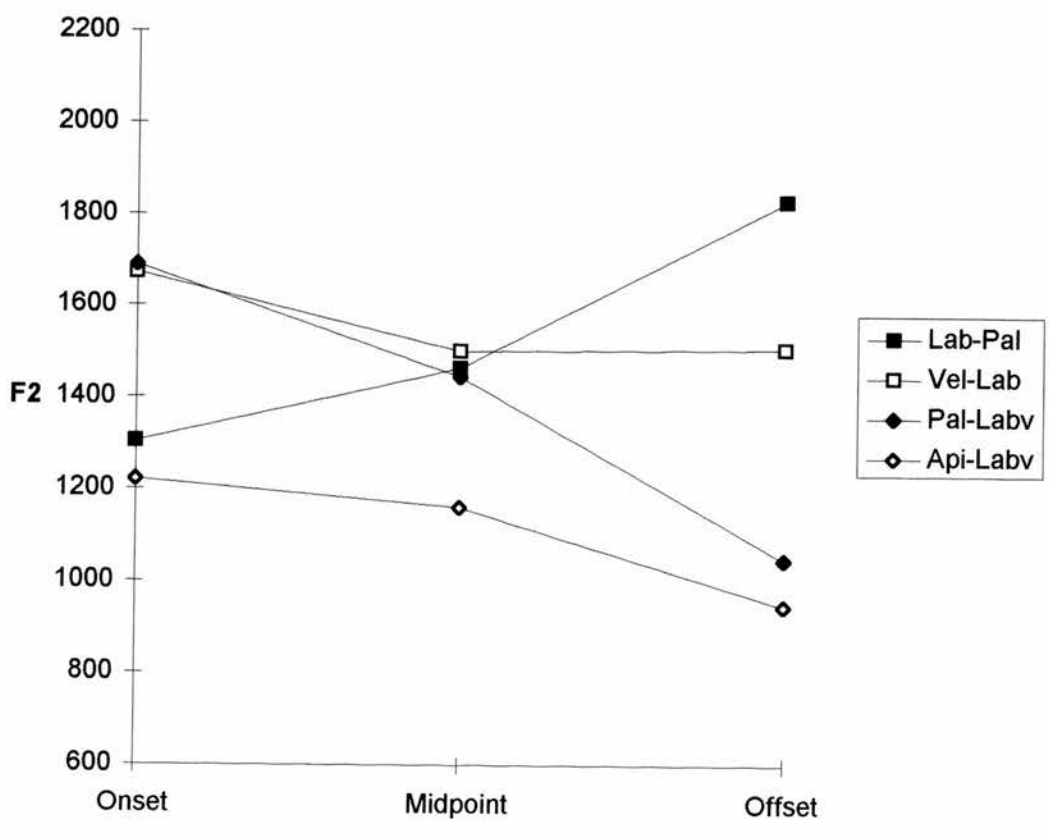
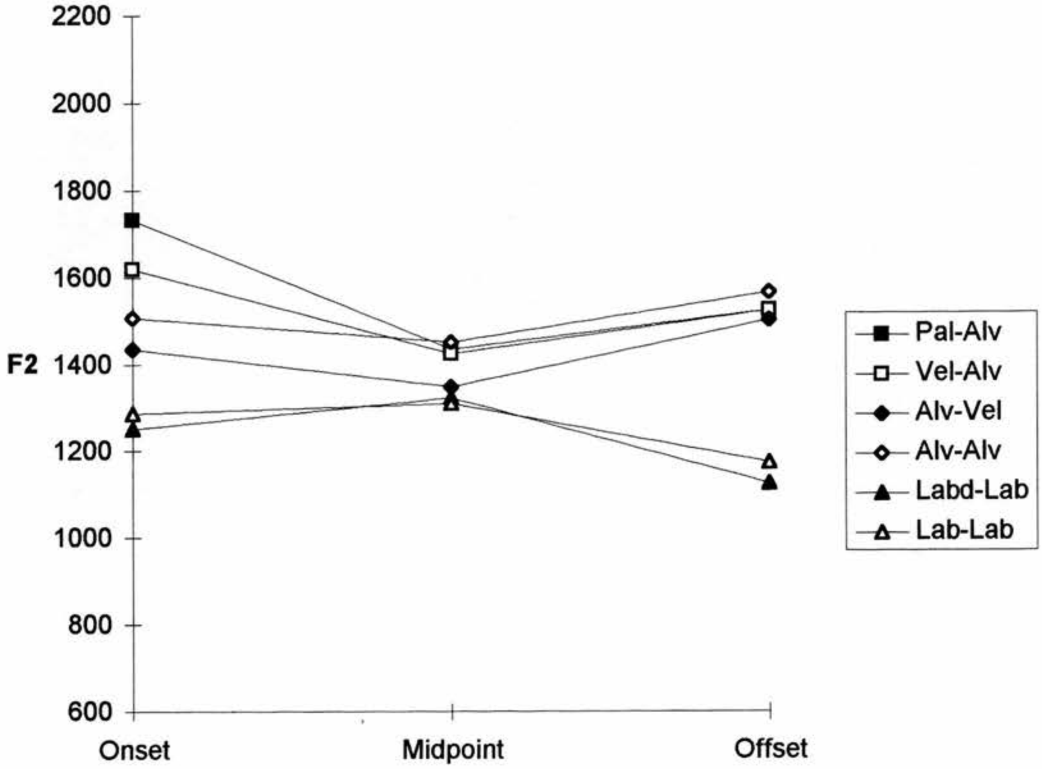


Figure 5.8: Mean second formant trajectories for /ɜ/

5.8a: high deviation values



5.8b: low deviation values

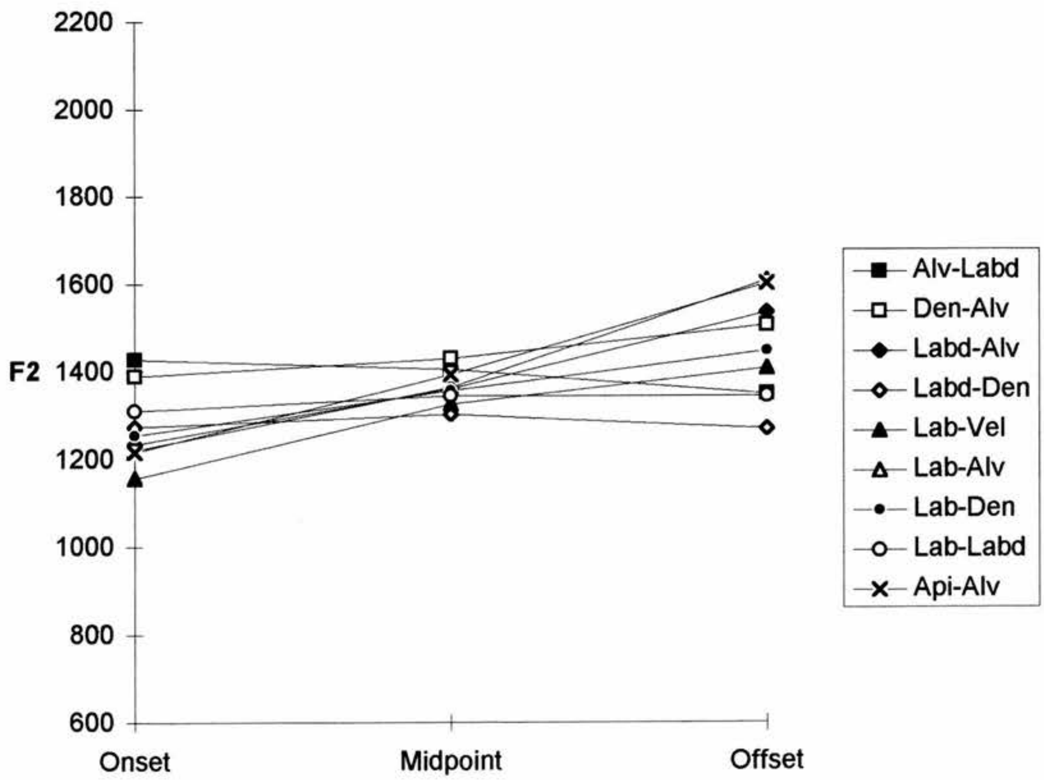
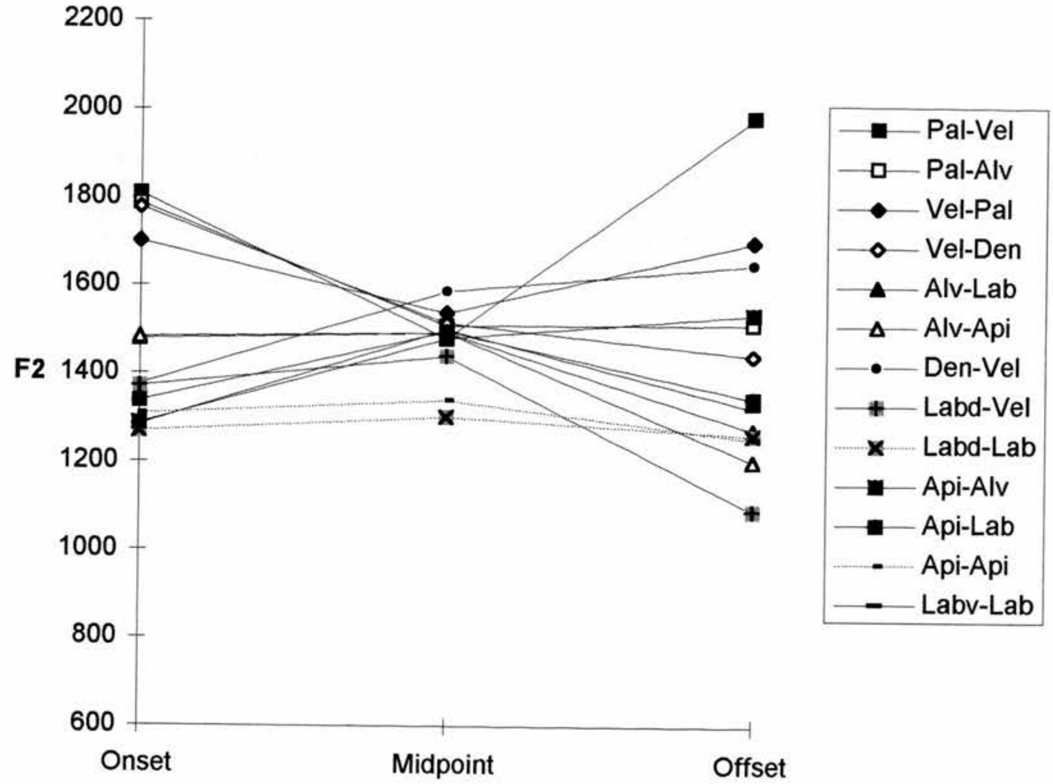


Figure 5.9: Mean second formant trajectories for /a/

5.9a: high deviation values



5.9b: low deviation values

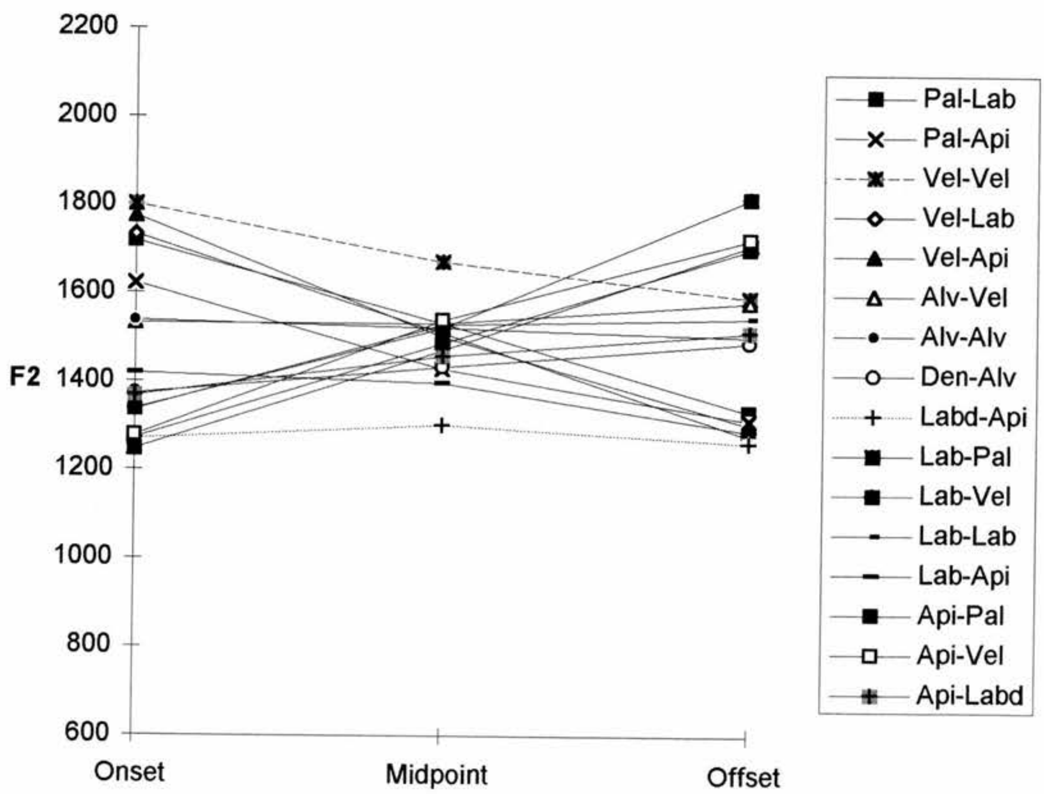


Figure 5.10: Mean second formant trajectories for /ə/, /ɪ/ and /ʊ/ for all contexts

5.10a: /ə/

5.10b: /ɪ/

5.10c: /ʊ/

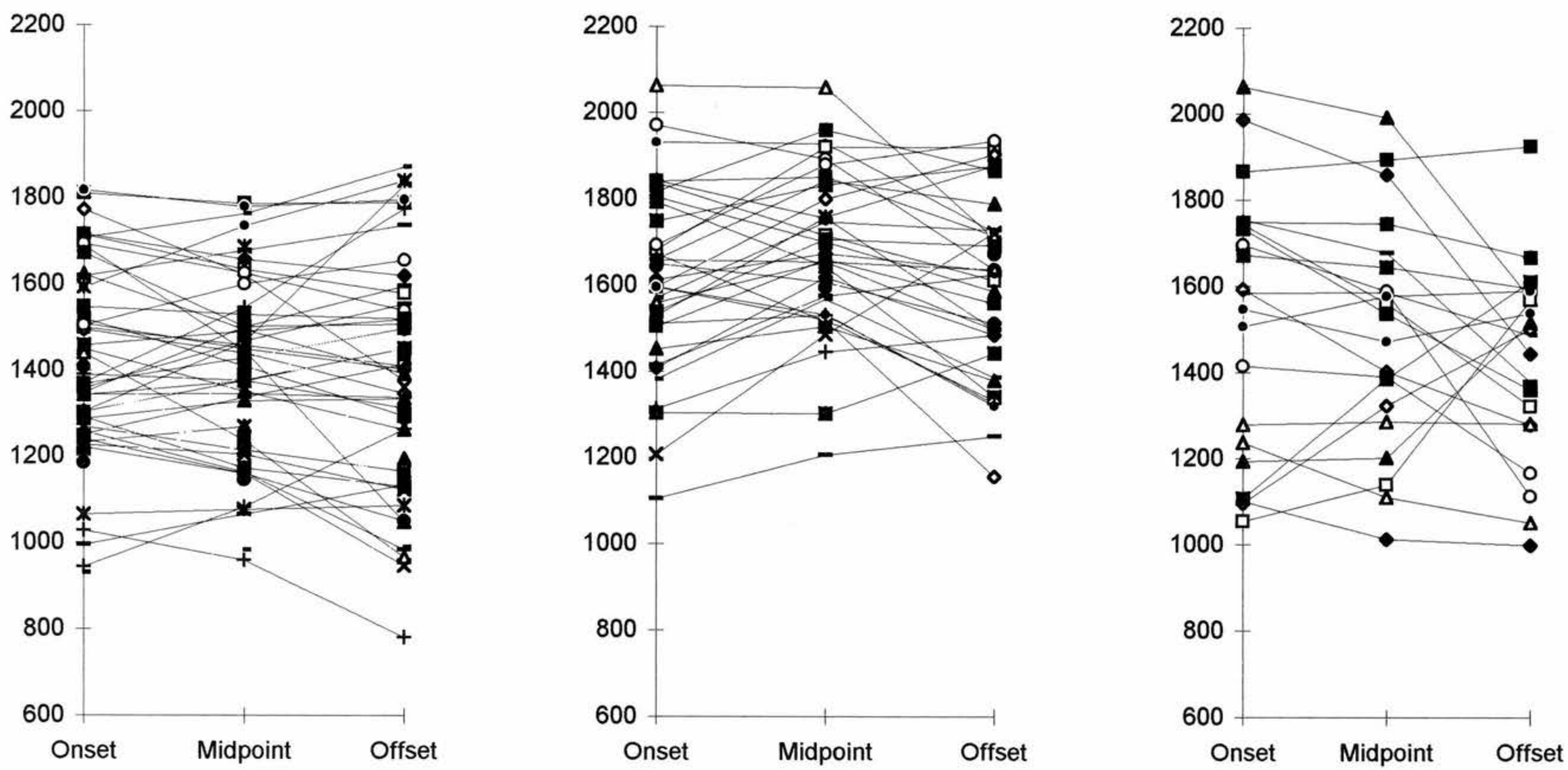


Figure 5.11: Mean second formant trajectories for /ɪ/

5.11a: "V-shaped" 5.11b: "L-shaped" 5.11c: "neo-L-shaped"

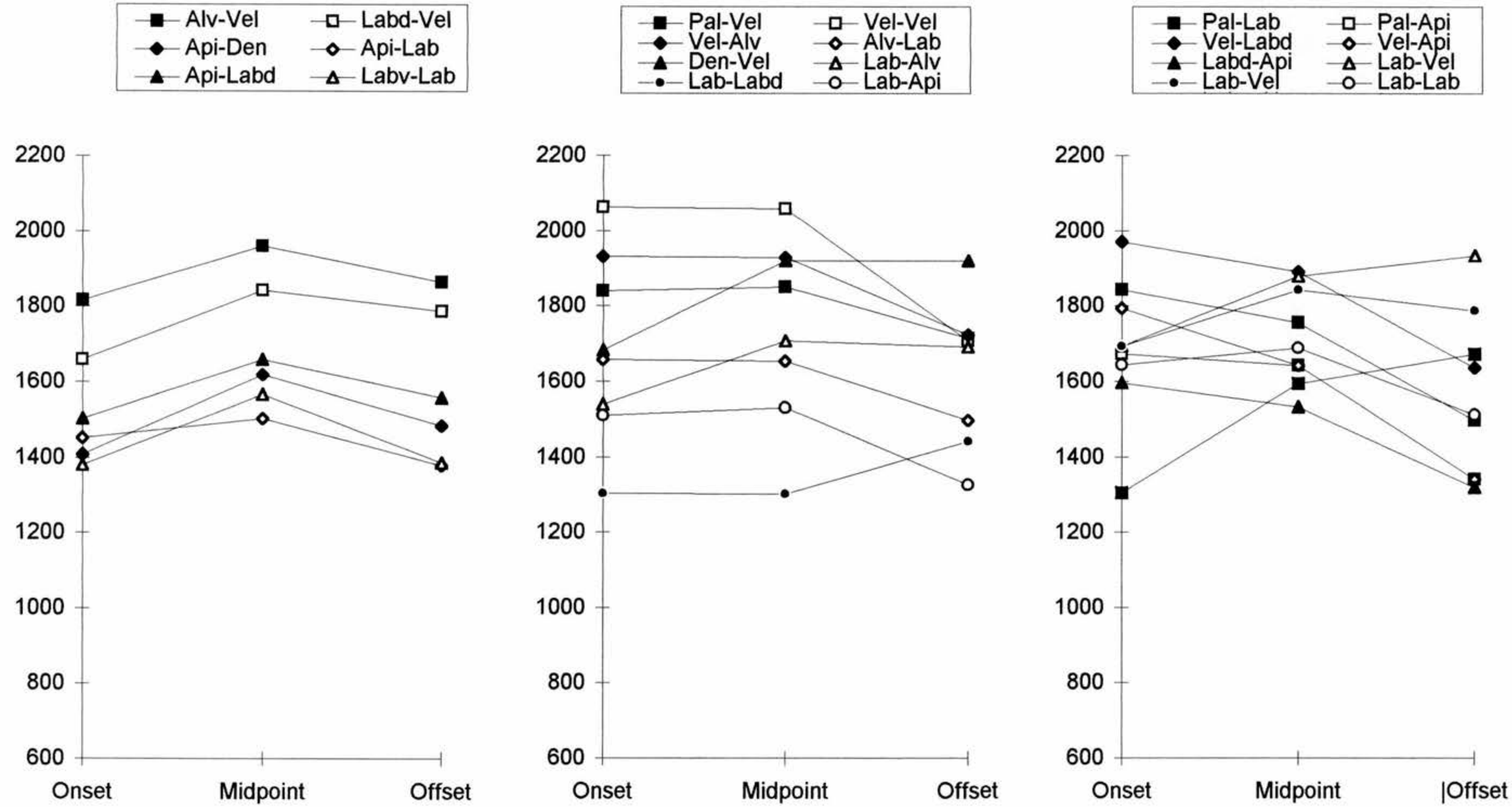
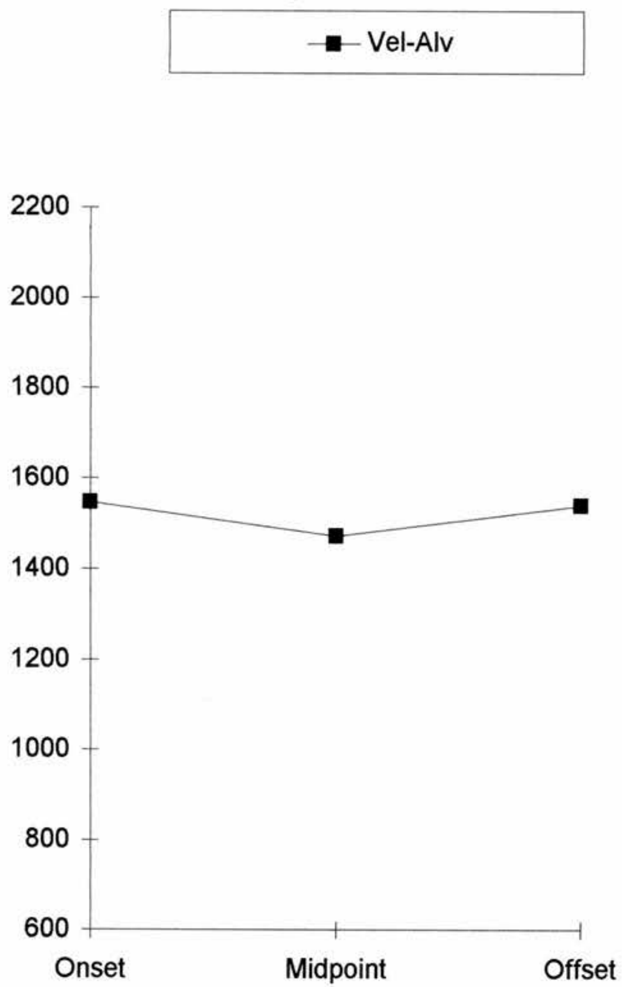
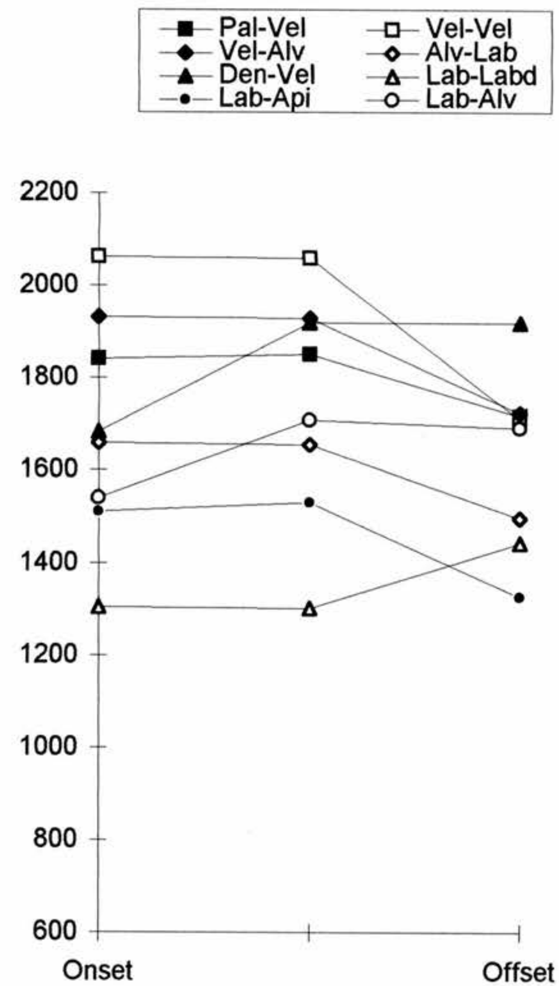


Figure 5.12: Mean second formant trajectories for /ʊ/

5.12a: "V-shaped"



5.12b: "L-shaped"



5.12c: "neo-L-shaped"

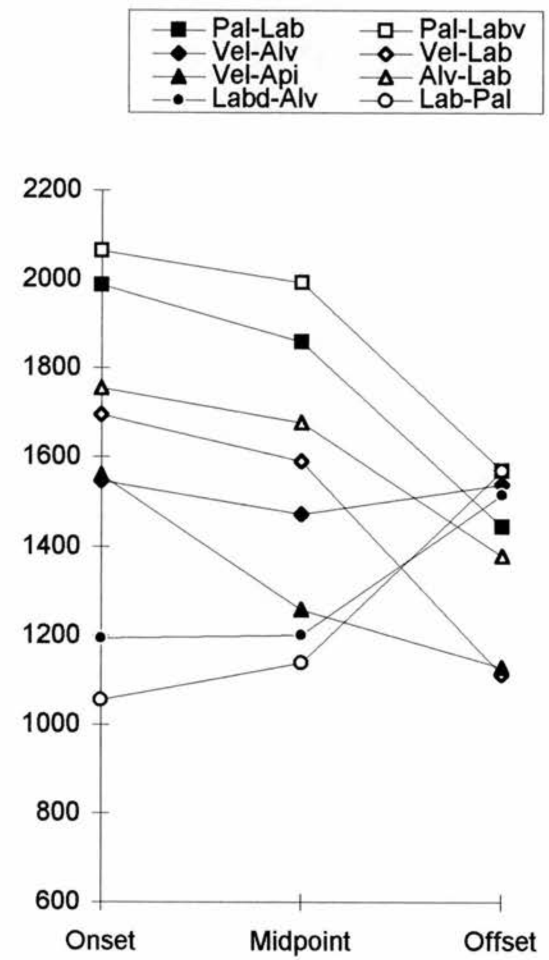
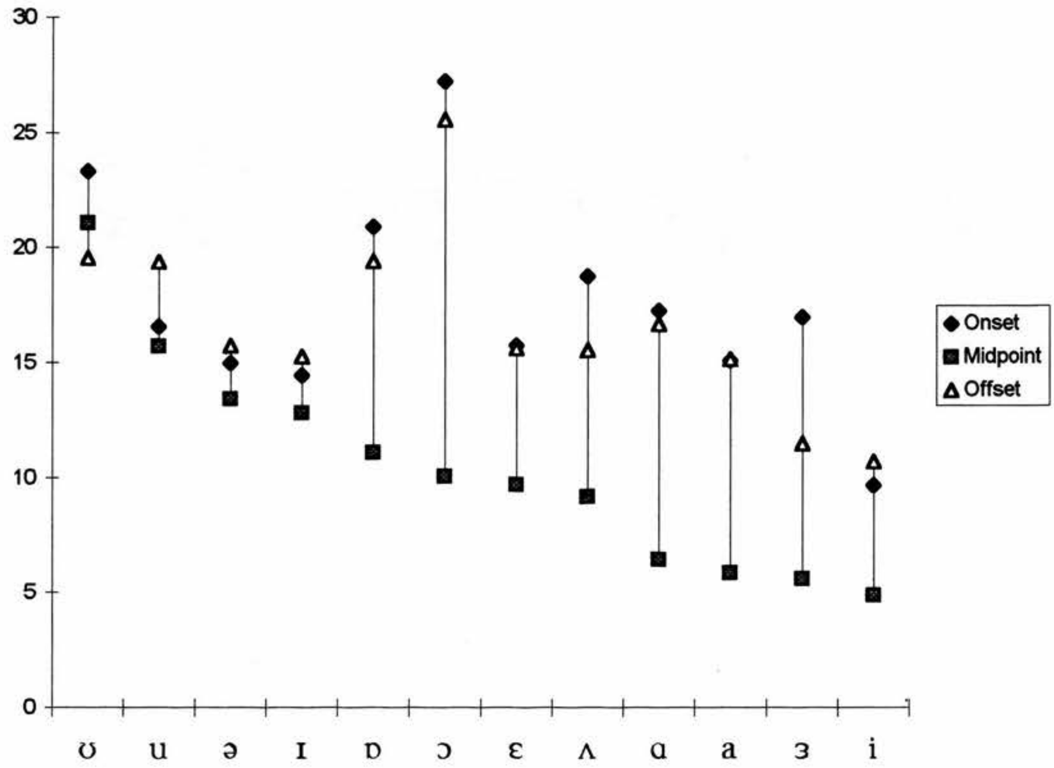


Figure 5.13: Coefficient of variance values

5.13a: F2



5.13b: F1

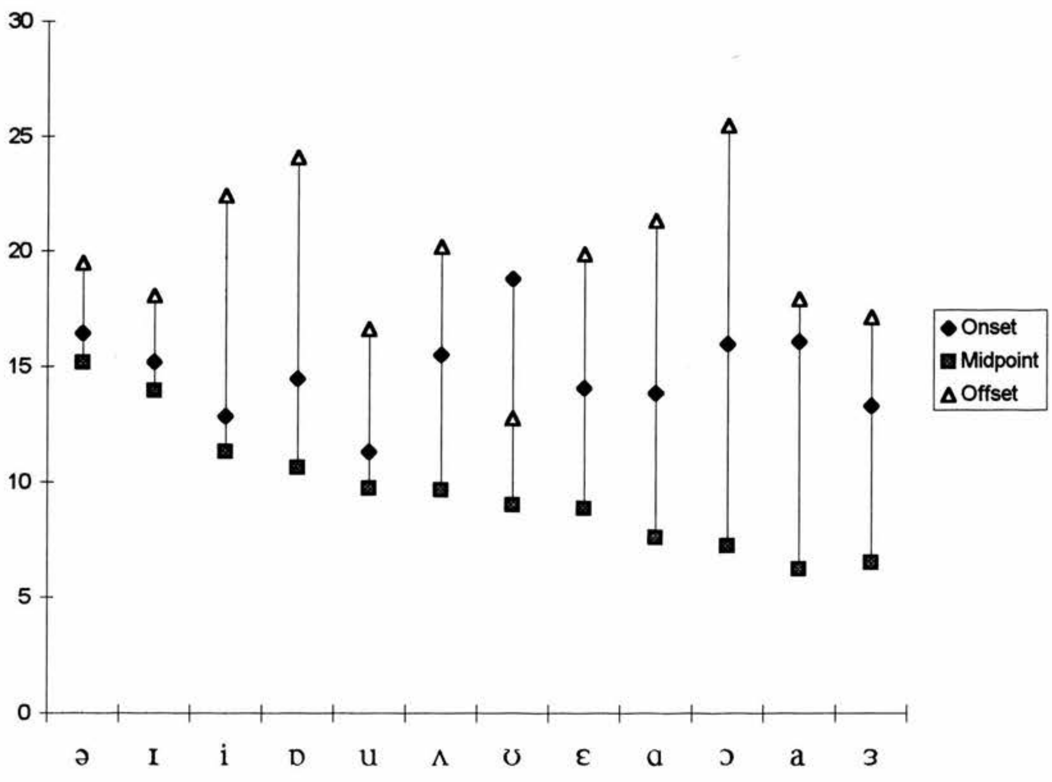
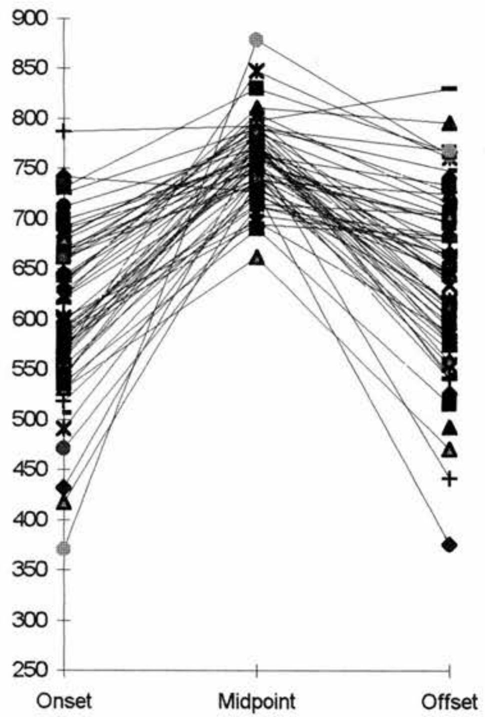
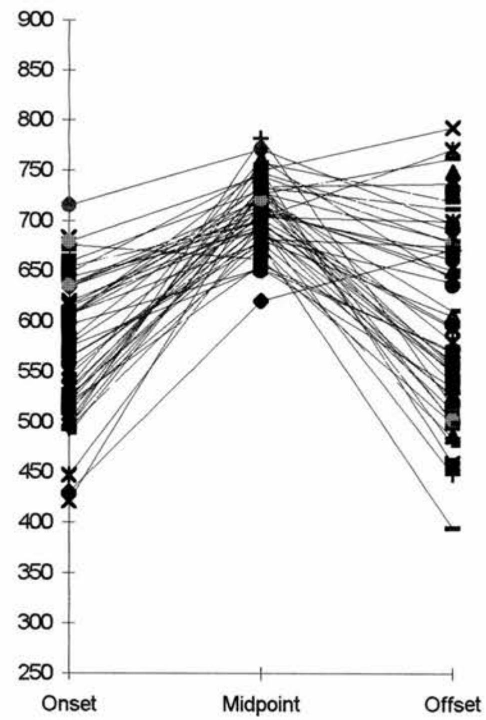


Figure 5.14: Mean first formant trajectories for /a/, /ɑ/ and /ʌ/ for all contexts

5.14a: /a/



5.14b: /ɑ/



5.14c: /ʌ/

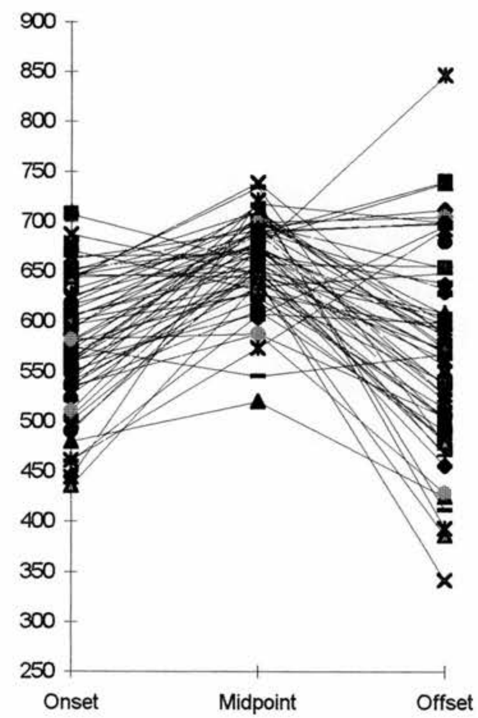
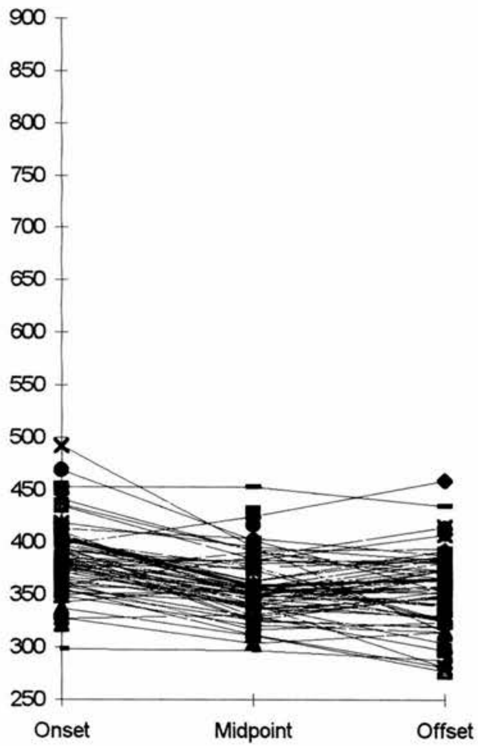
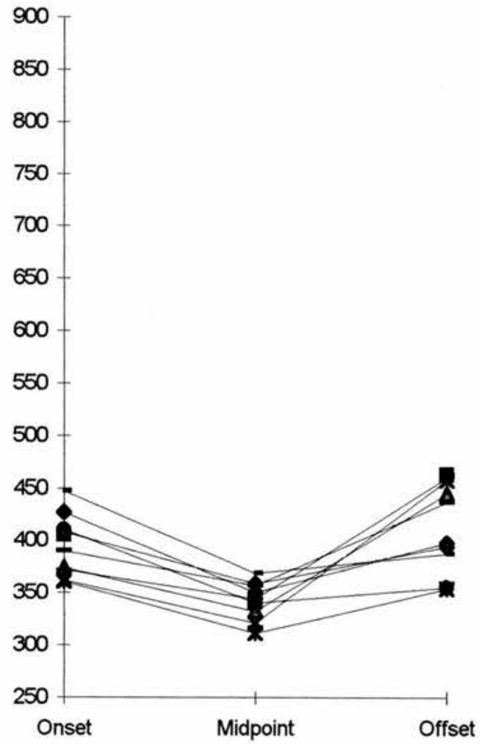


Figure 5.15: Mean first formant trajectories for /i/

5.15a: low deviation values



5.15b: high deviation values - "concave downward"



5.15c: high deviation values - "concave upward"

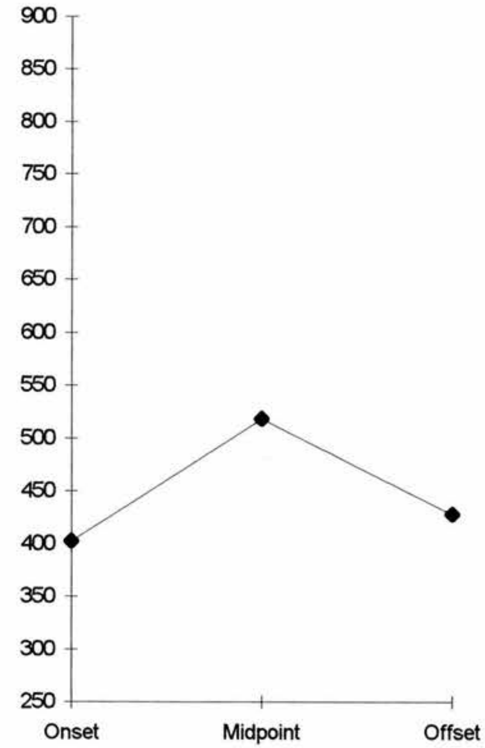
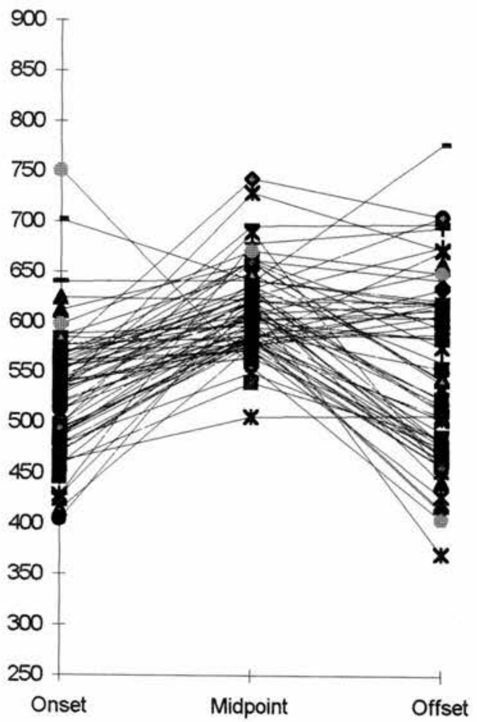
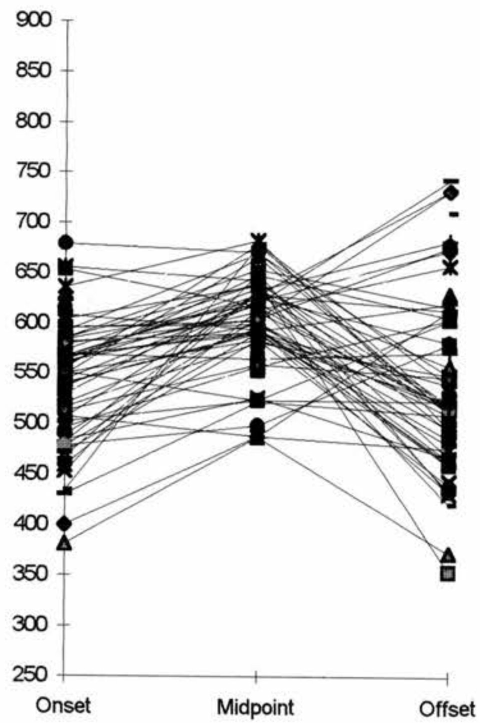


Figure 5.16: Mean first formant trajectories for /ɛ/, /ɒ/ and /ɔ/ for all contexts

5.16a: /ɛ/



5.16b: /ɒ/



5.16c: /ɔ/

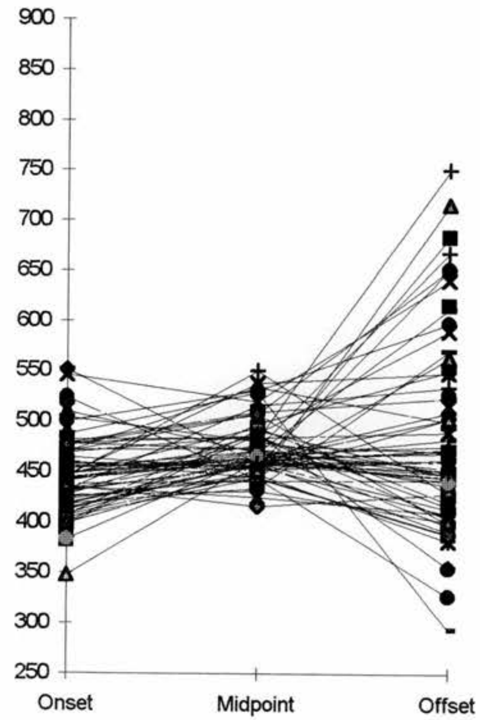
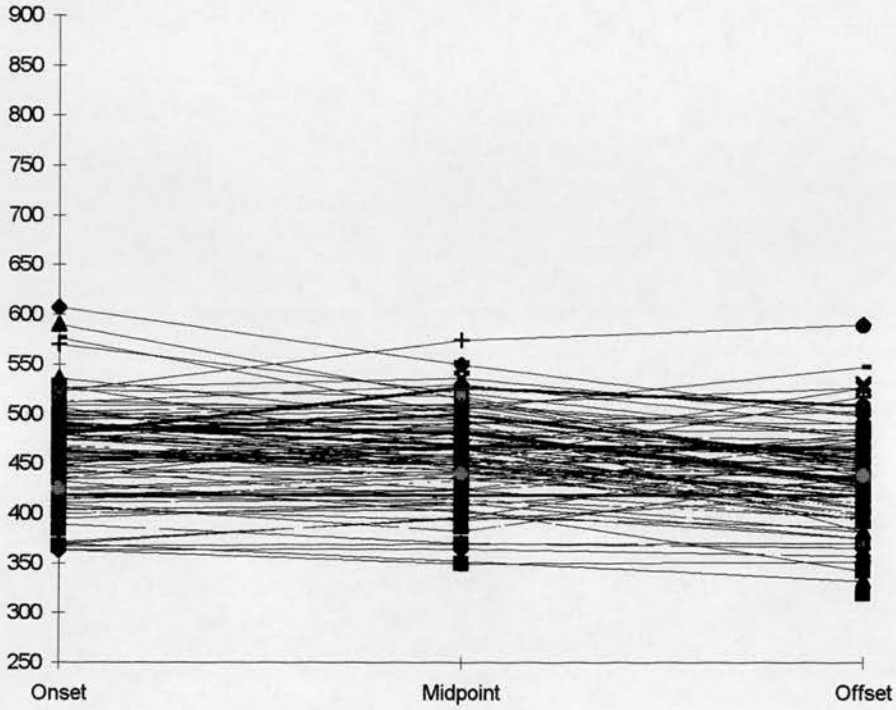


Figure 5.17: Mean first formant trajectories for schwa

5.17a: low deviation values



5.17b: high deviation values

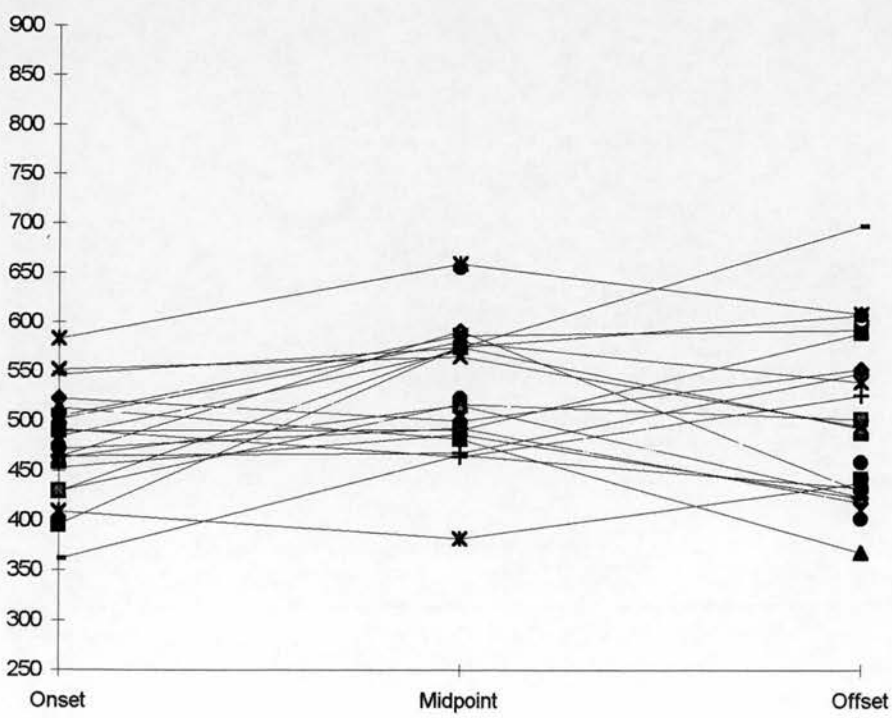


Figure 5.18: Mean first formant trajectories for schwa - high deviation values

5.18a: "L-shaped" 5.18b: "V-shaped" - concave upward" 5.18c:"V-shaped" - concave downward

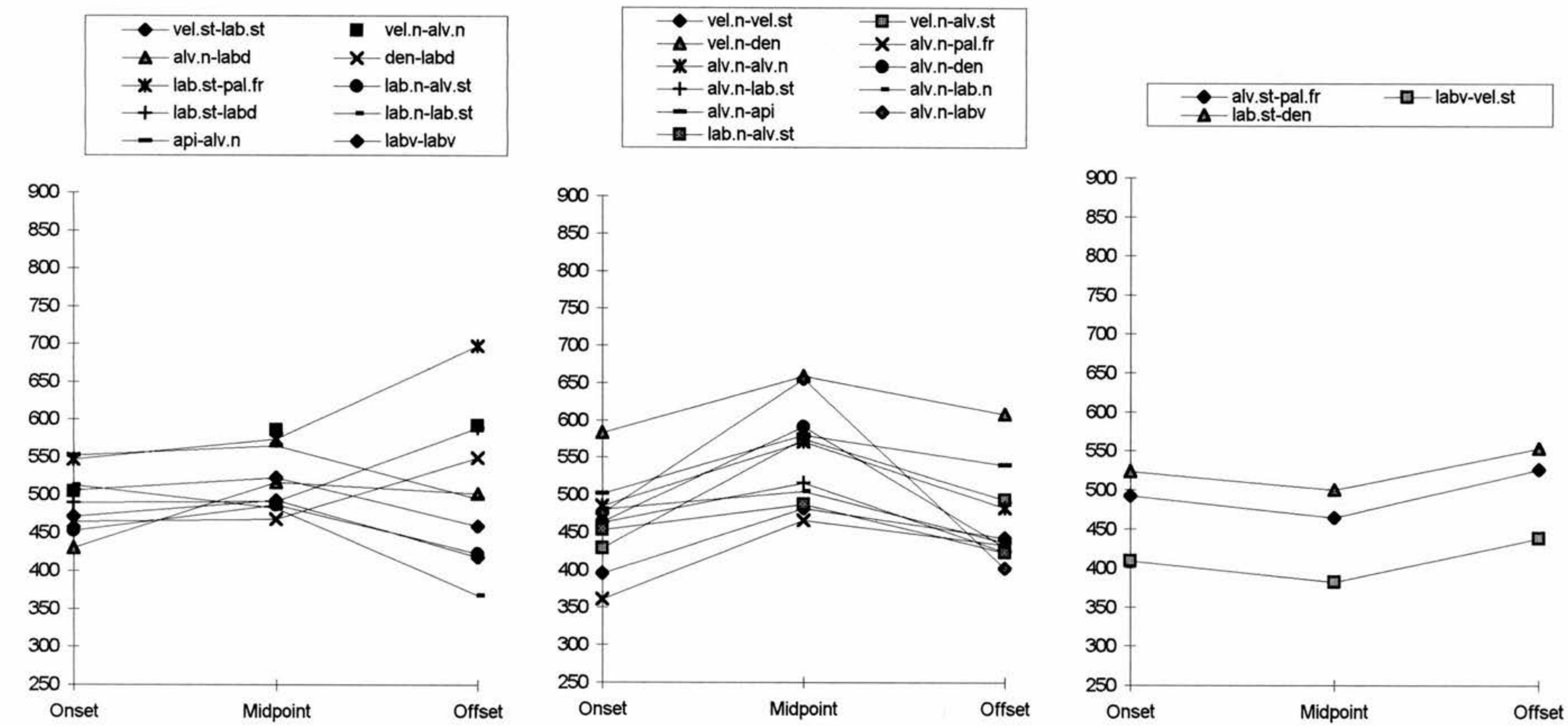


Figure 5.19: Mean first formant trajectories for schwa for all contexts

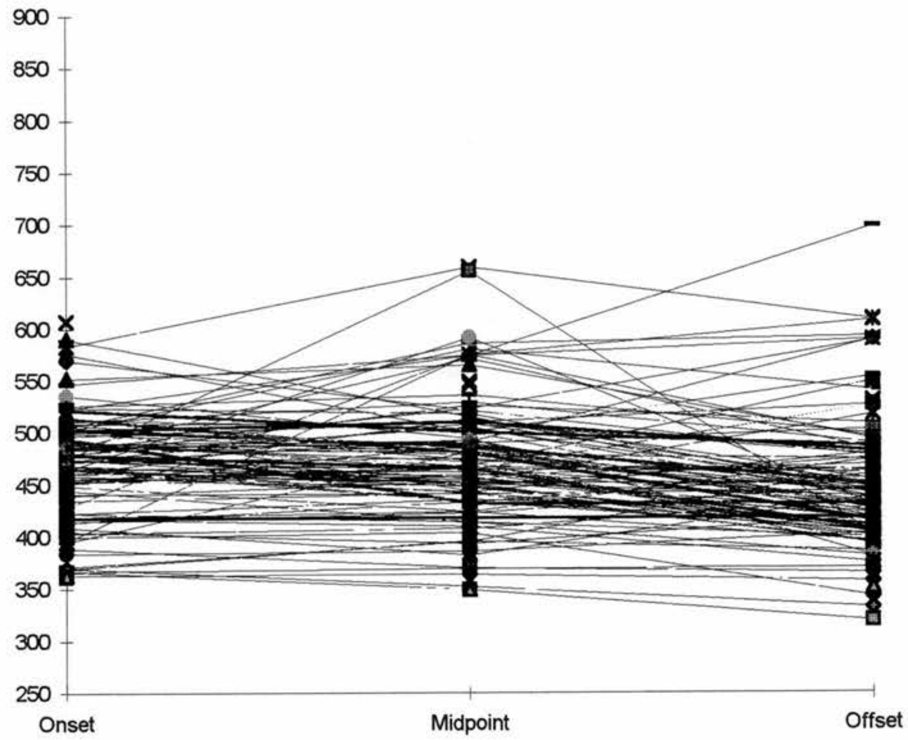


Figure 5.20: Mean first formant trajectories for /ɜ/ for all contexts

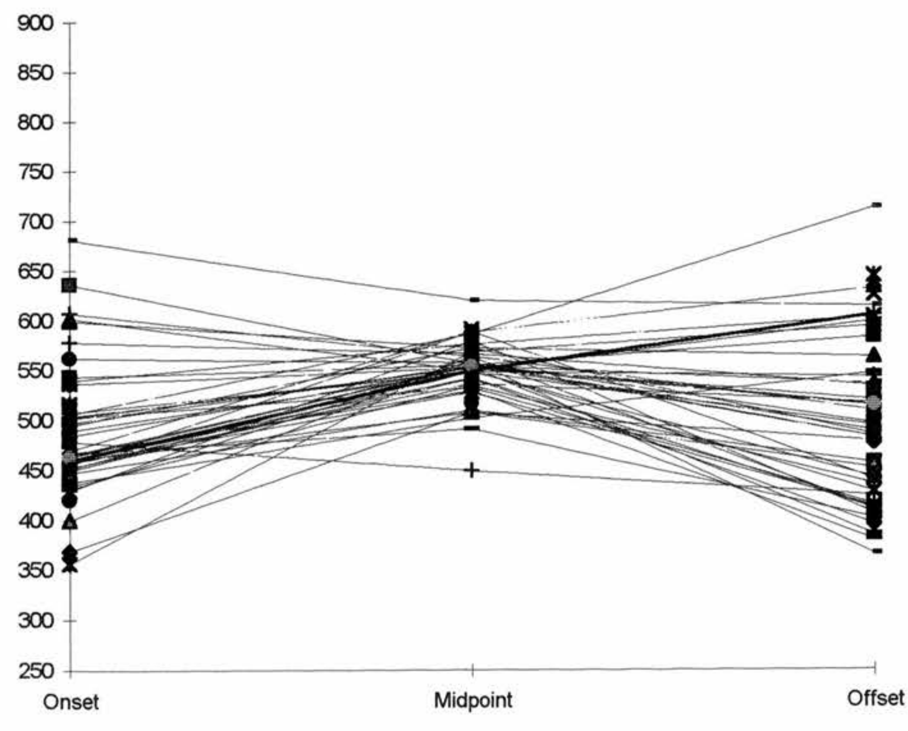


Figure 5.21: Mean first formant trajectories for /ɜ/ - low deviation values

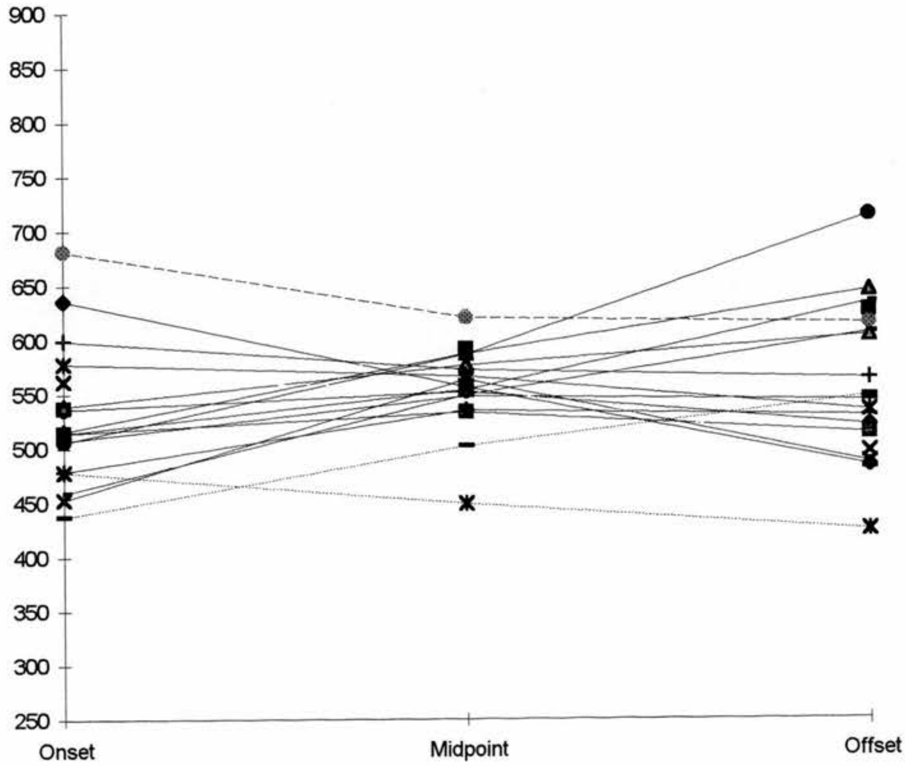


Figure 5.22: Mean first formant trajectories for schwa - low deviation values

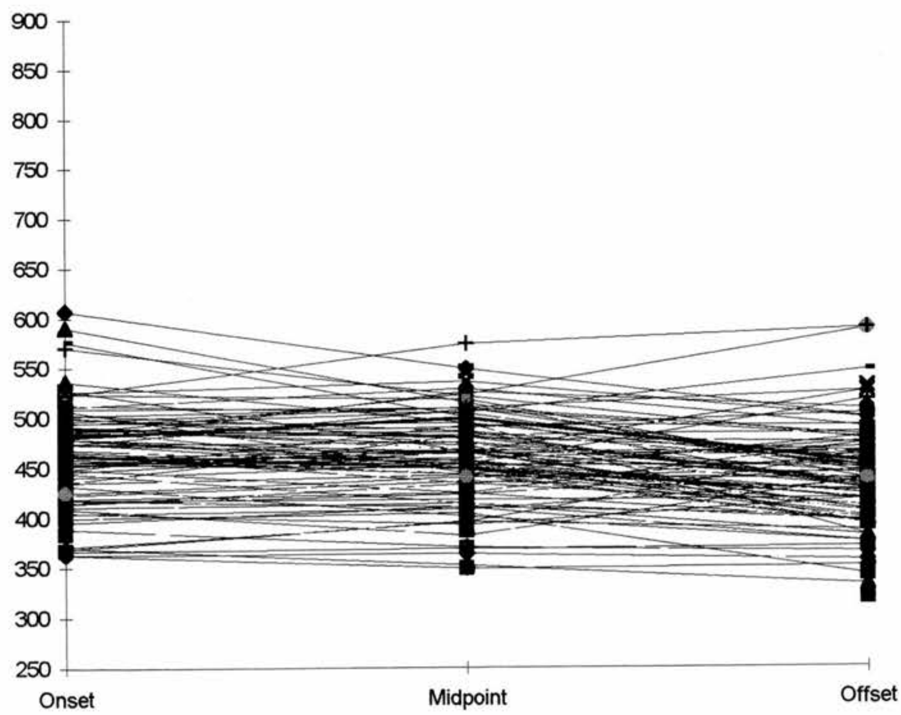


Figure 5.23: Mean first formant trajectories for /ɜ/ - high deviation values

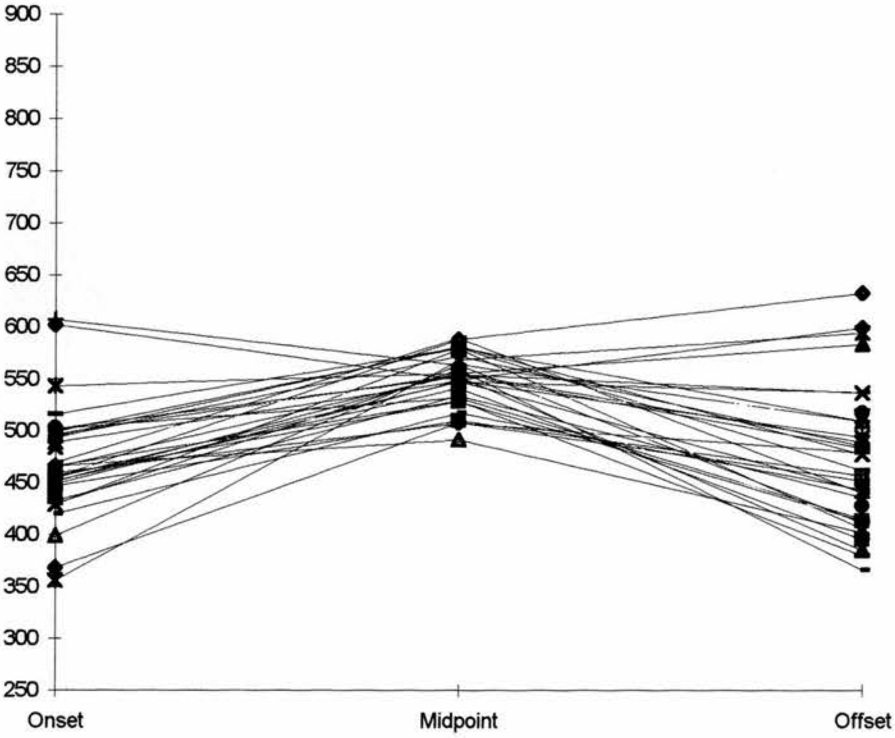


Figure 5.24: Mean first formant trajectories for schwa - high deviation values

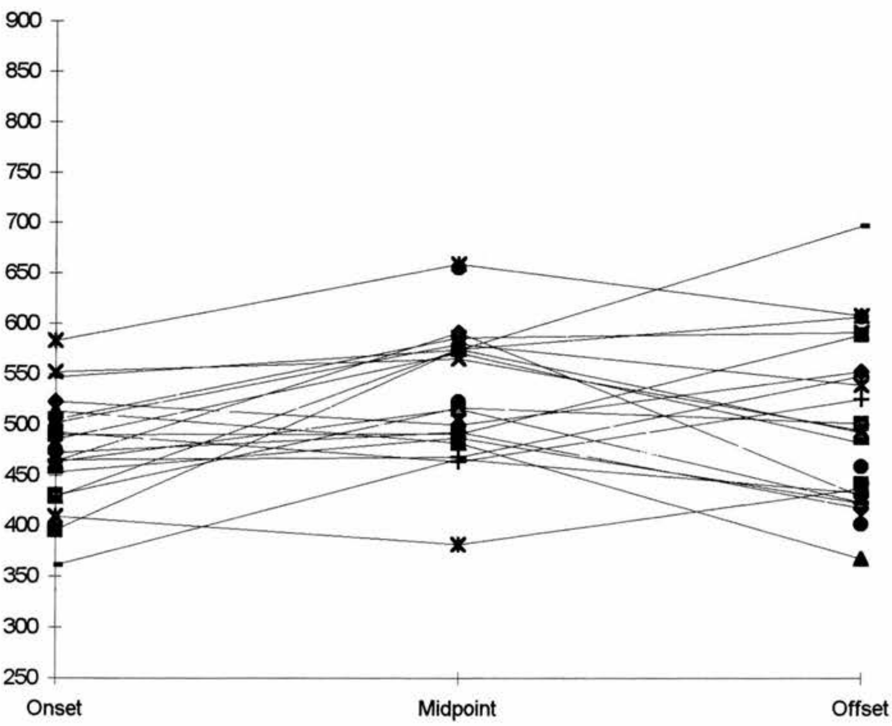


Figure 5.25: Mean first formant trajectories for /ə/, /ɪ/ and /ʊ/ for all contexts

5.25a: /ə/ 5.25b: /ɪ/ 5.25c: /ʊ/

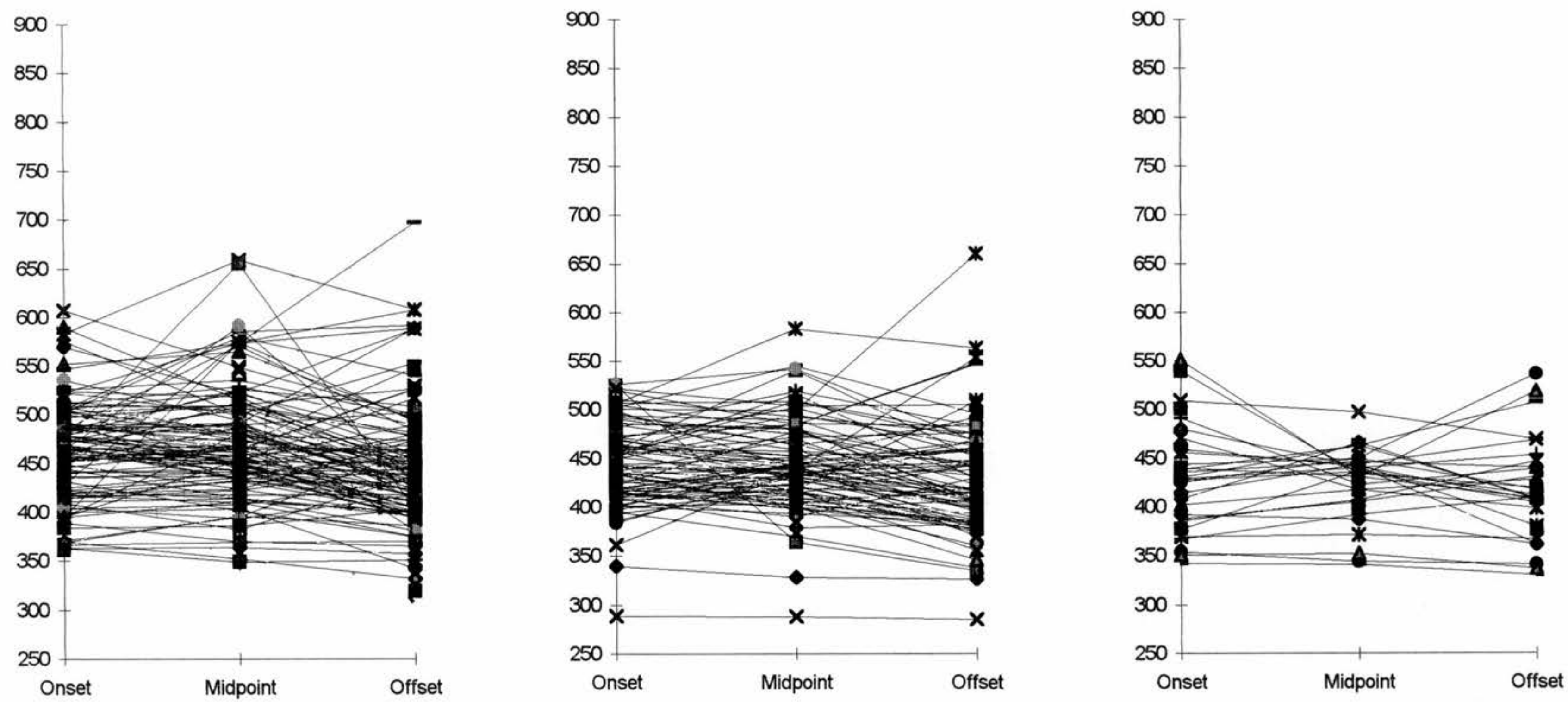
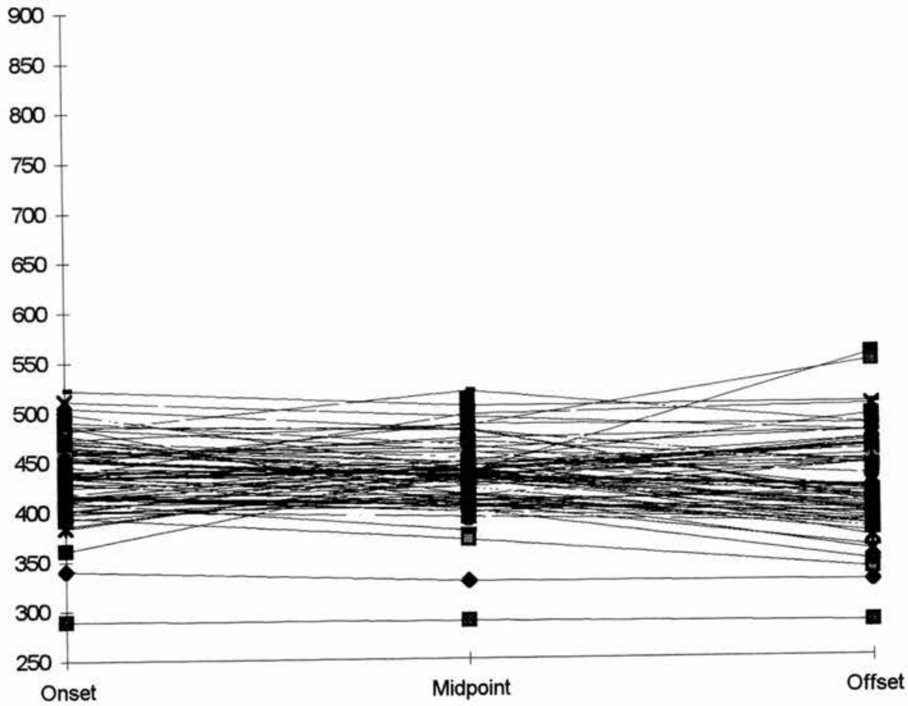


Figure 5.26: Mean first formant trajectories for /ɪ/

5.26a: low deviation values



5.26b: high deviation values

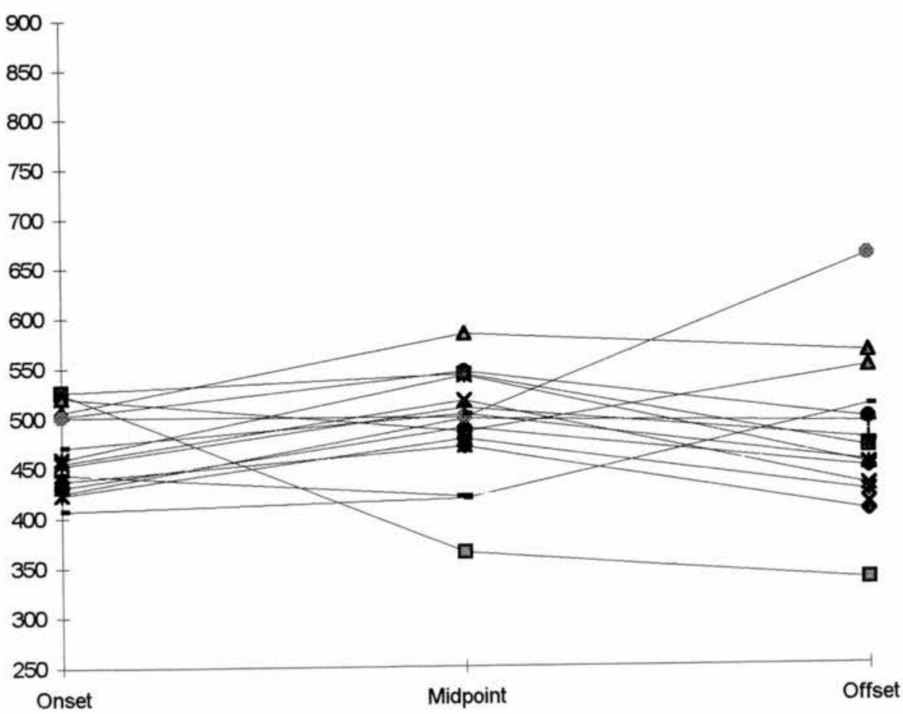
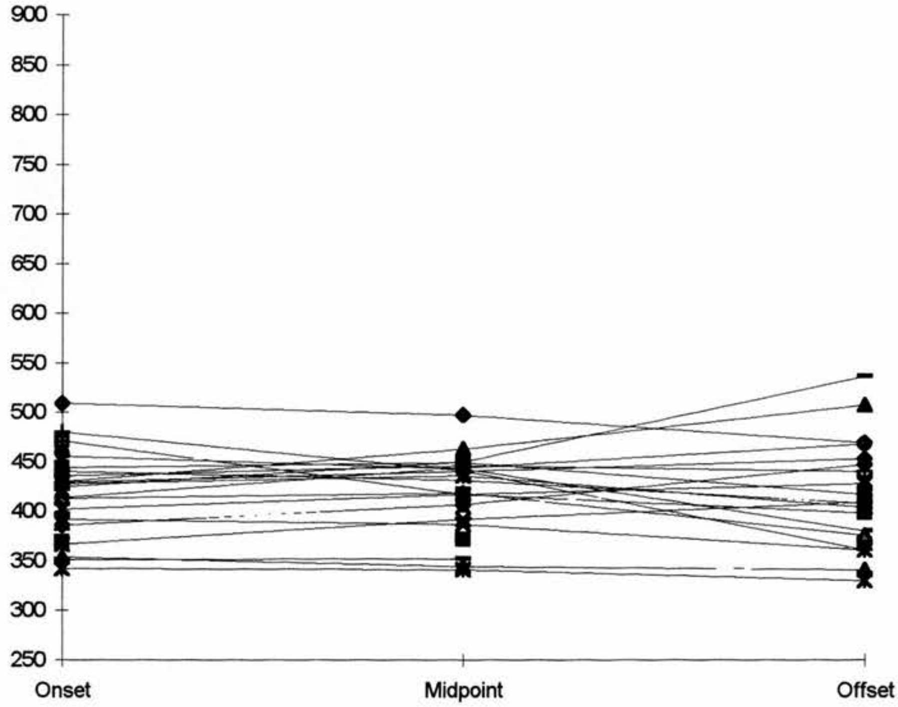


Figure 5.27: Mean first formant trajectories for /ʊ/

5.27a: low deviation values



5.27b: high deviation values

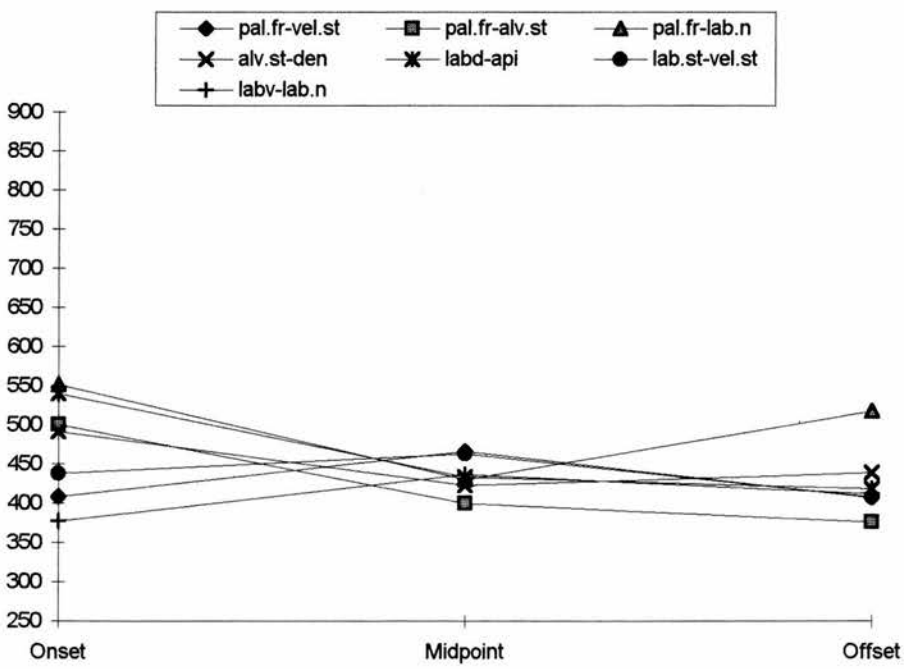
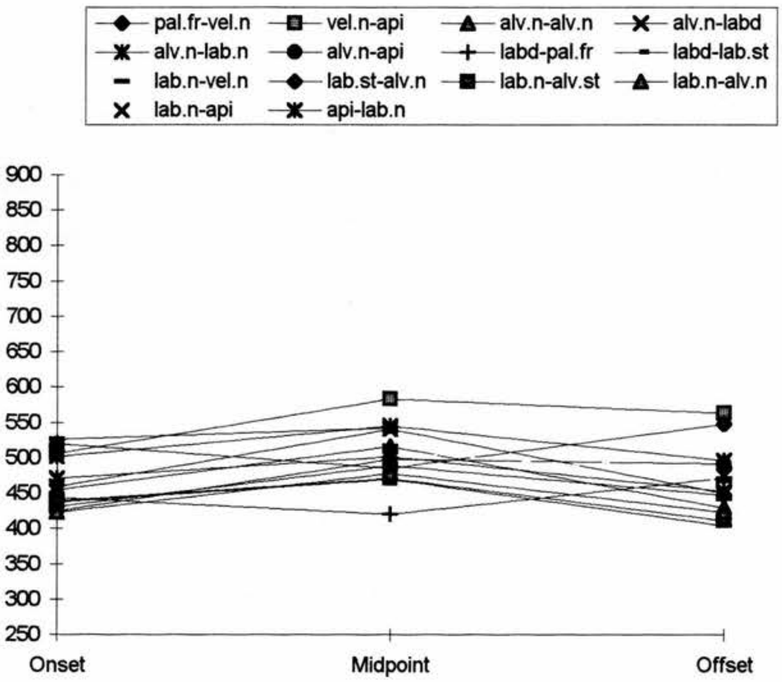


Figure 5.28: Mean first formant trajectories for /ɪ/ - high deviation values

5.28a: "V-shaped"



5.28b: "L-shaped"

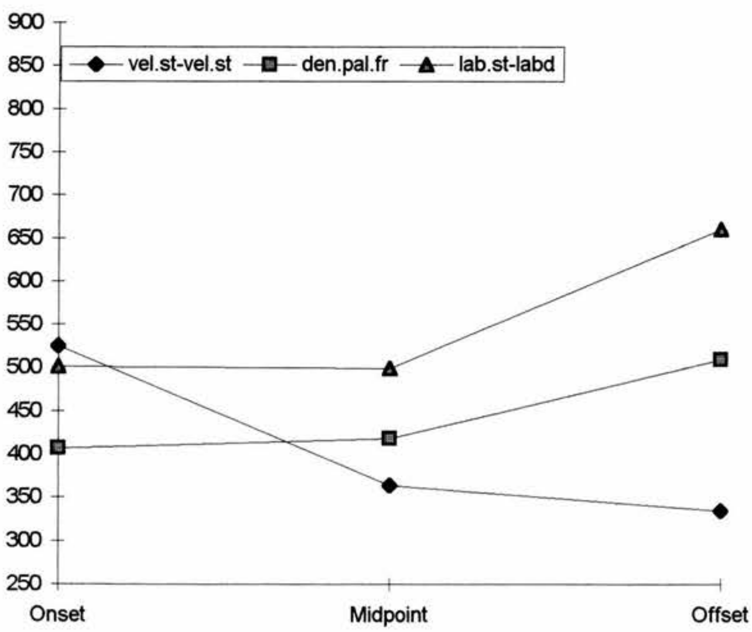


Figure 5.29: Second formant trajectories for schwa.
F2 is plotted along the vertical axis. The horizontal axis represents normalised duration. The interval between 0 and 1 corresponds to the schwa vocalic interval. The mrpa symbols /i/ and /oo/ denote /i/ and /ɔ/ respectively.

5.29a: preceding context = /i/

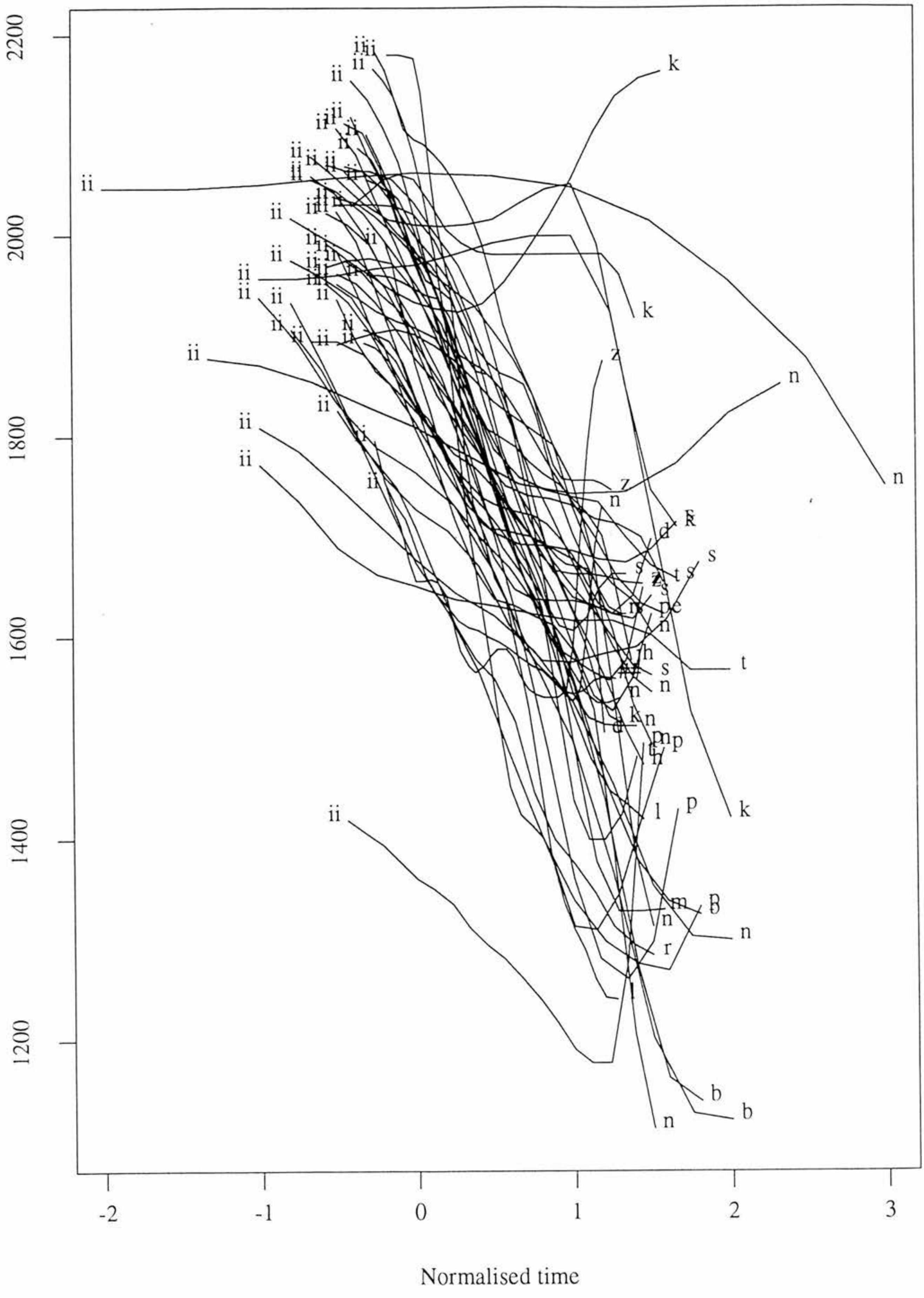


Figure 5.29: Second formant trajectories for schwa (continued).

5.29b: preceding context = /ɔ/

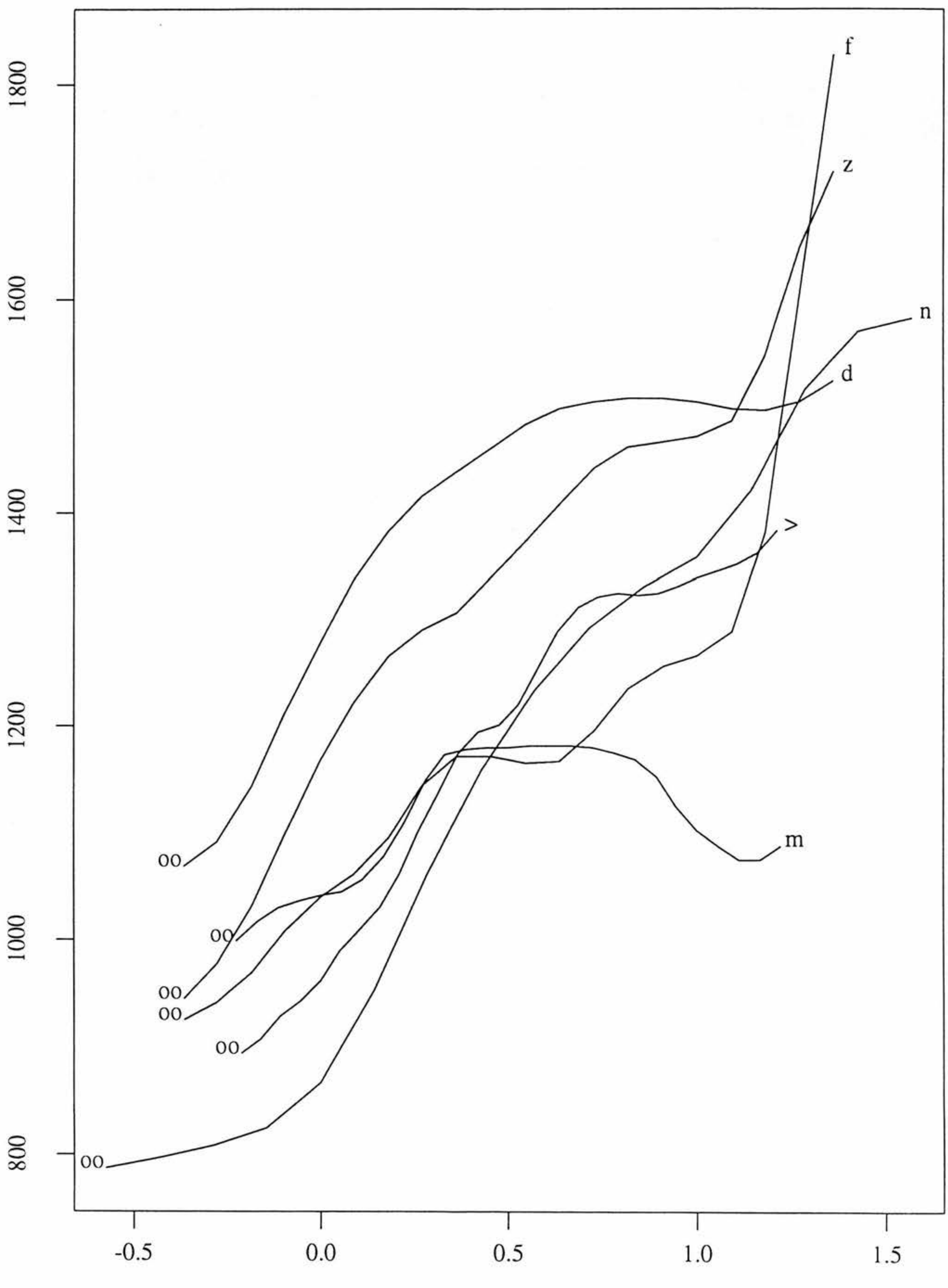


Figure 5.30: Scatterplot of schwa F2 onset versus midpoint values
Onset values are plotted along the y-axis and midpoint values along the x-axis.

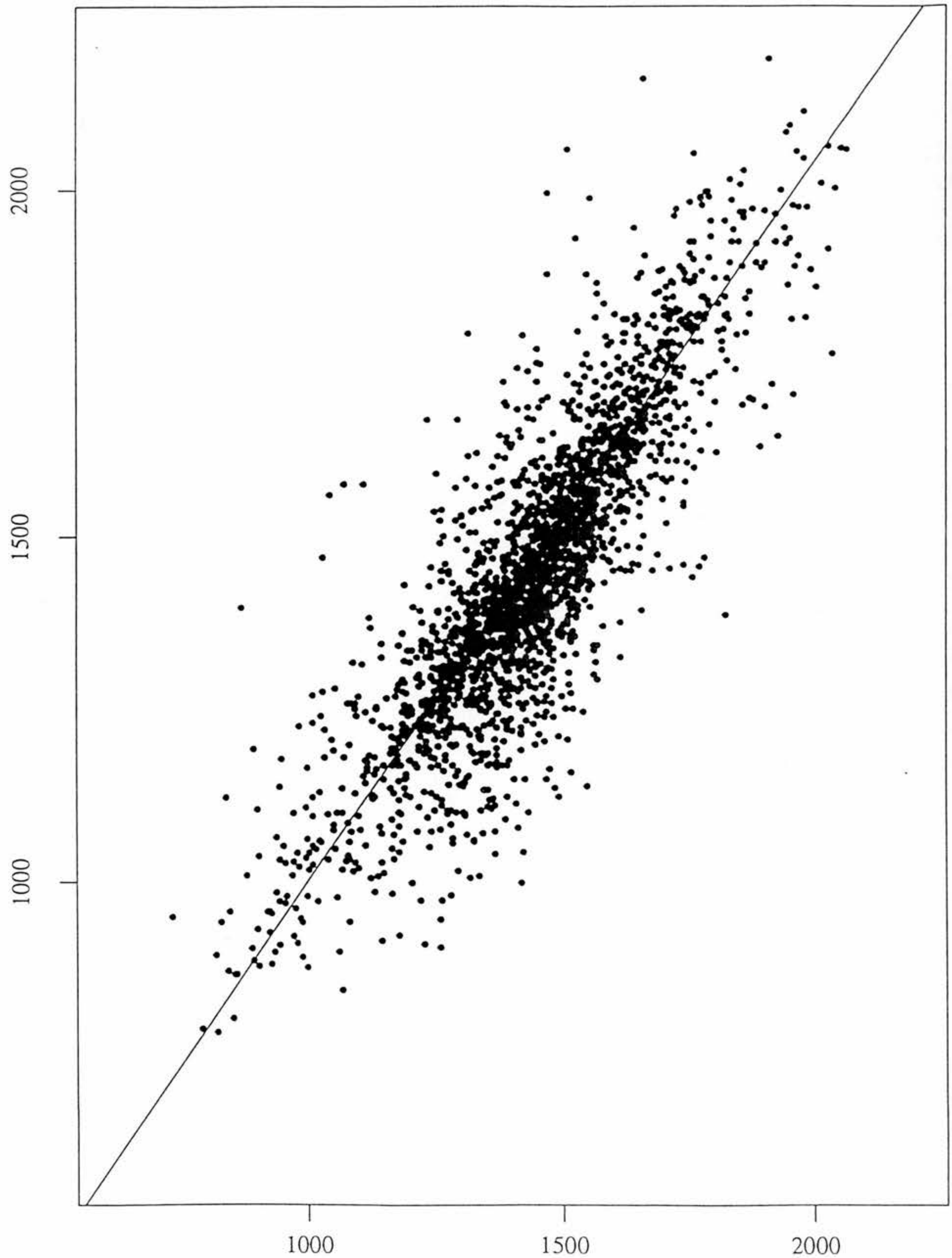


Figure 5.31: Scatterplot of /ɪ/ F2 onset versus midpoint values
Onset values are plotted along the y-axis and midpoint values along the x-axis.

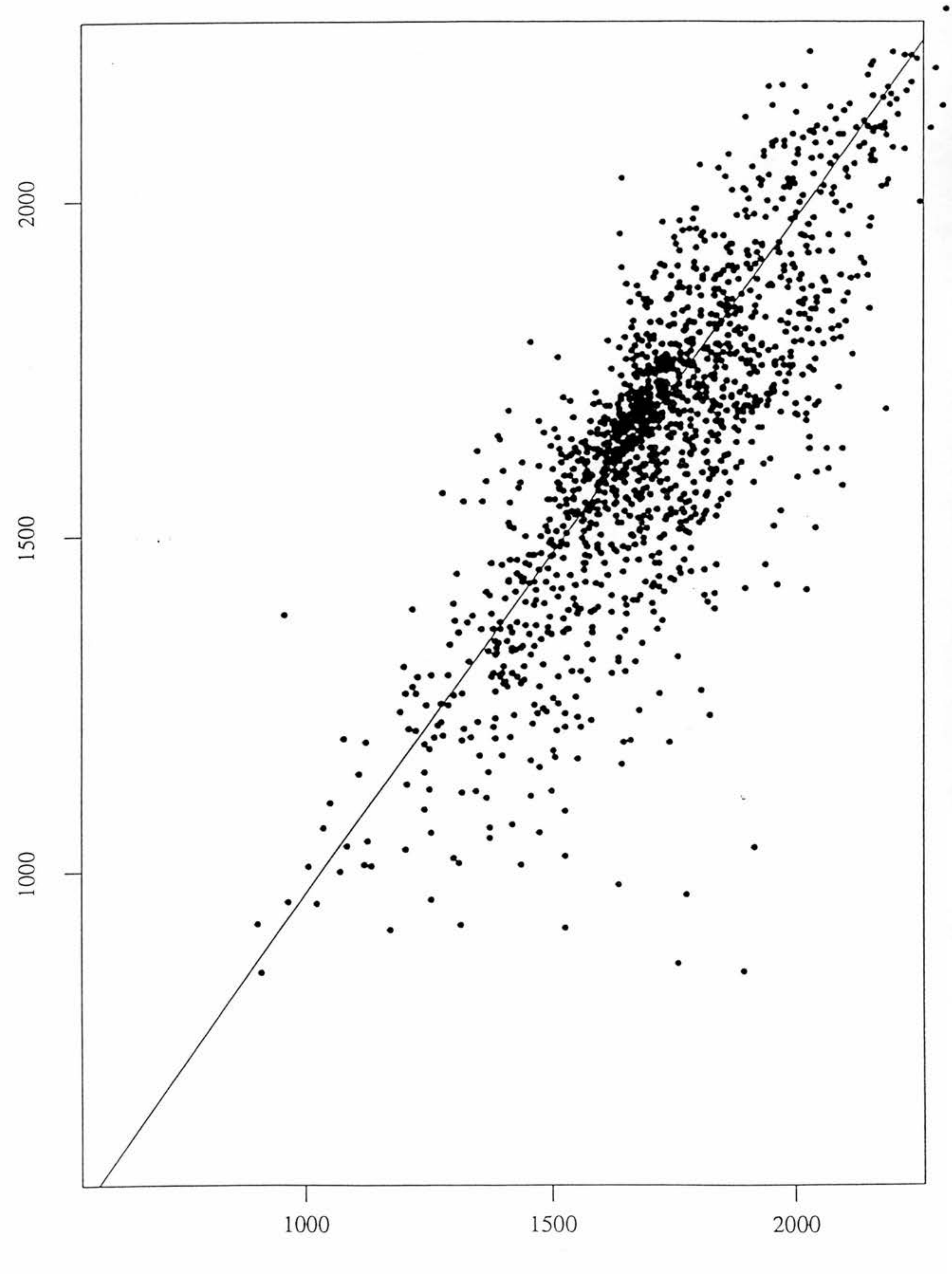


Figure 5.32: Scatterplot of /ʊ/ F2 onset versus midpoint values
Onset values are plotted along the y-axis and midpoint values along the x-axis.

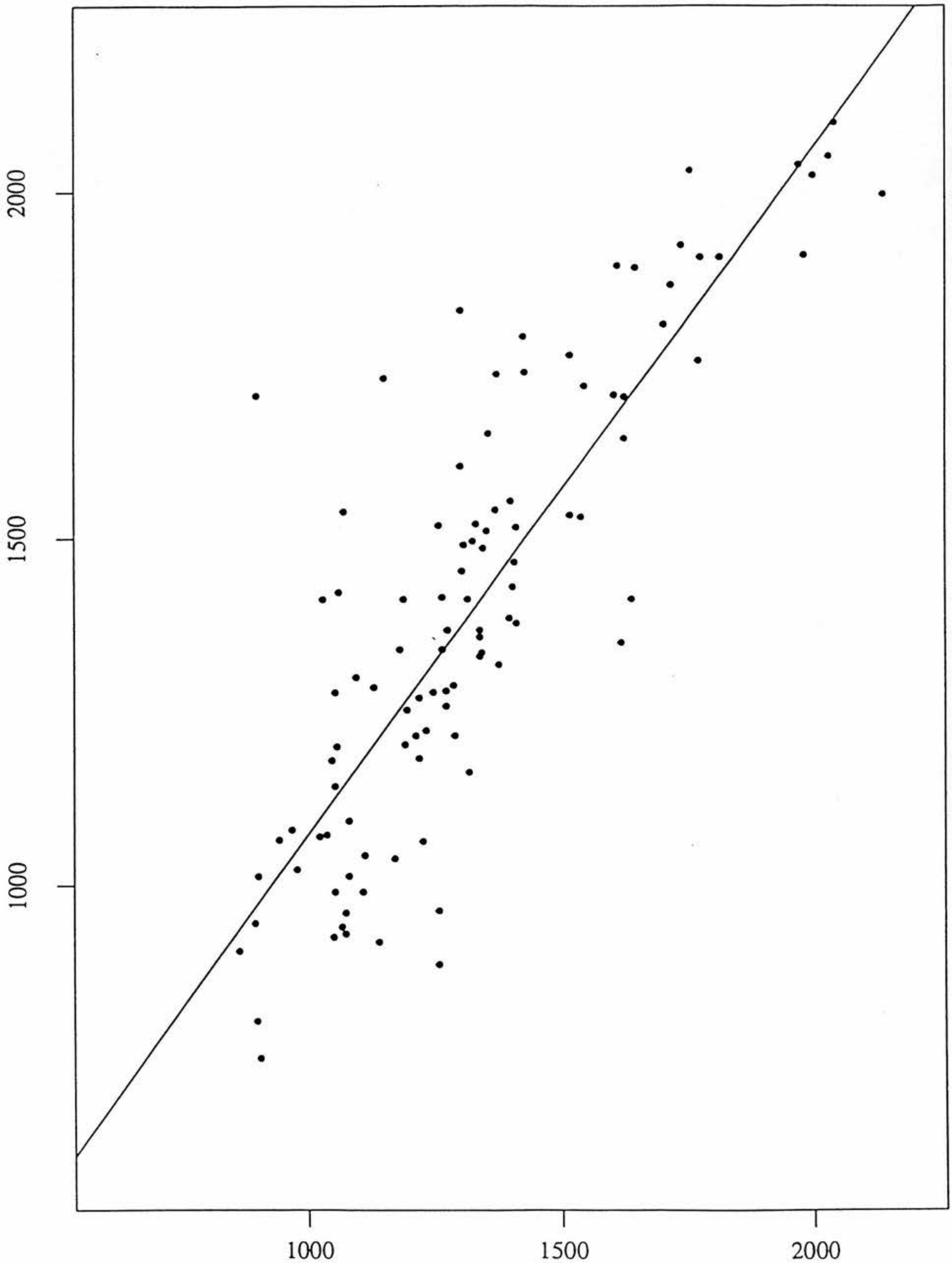


Figure 5.33: Scatterplot of /i/ F2 onset versus midpoint values
Onset values are plotted along the y-axis and midpoint values along the x-axis.

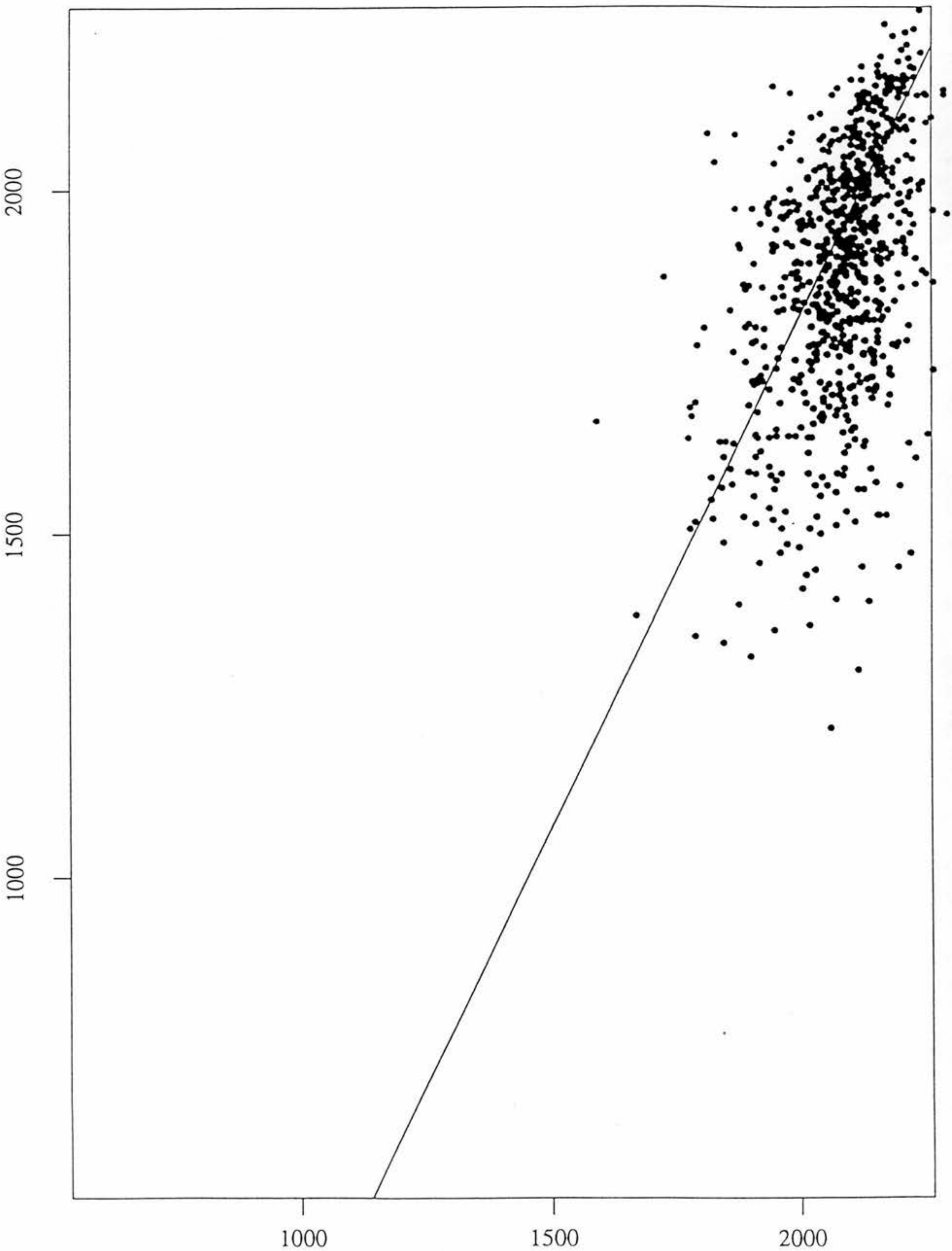


Figure 5.34: Scatterplot of /ɔ/ F2 onset versus midpoint values
Onset values are plotted along the y-axis and midpoint values along the x-axis.

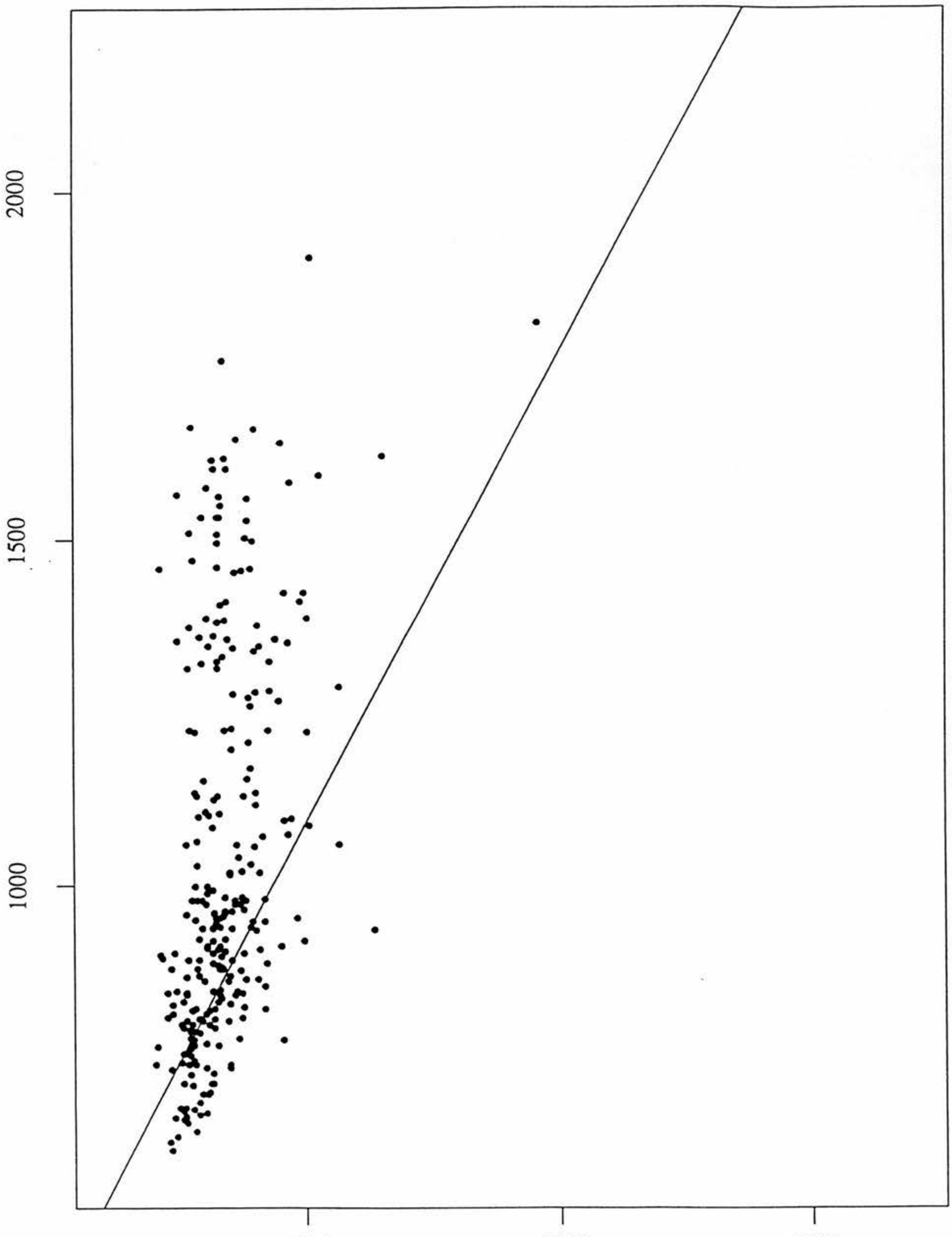
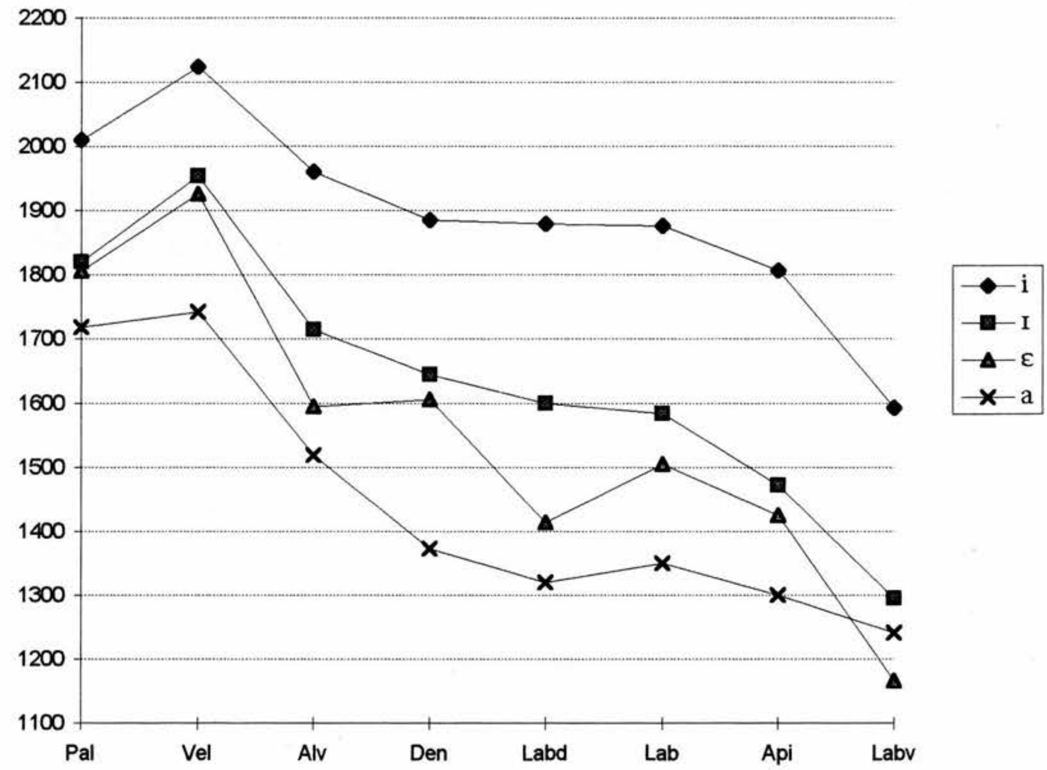


Figure 5.36: F2 mean onset/offset values as a function of C^1/C^2 place of articulation - front vowels

5.36a: mean onset values



5.36b: mean offset values

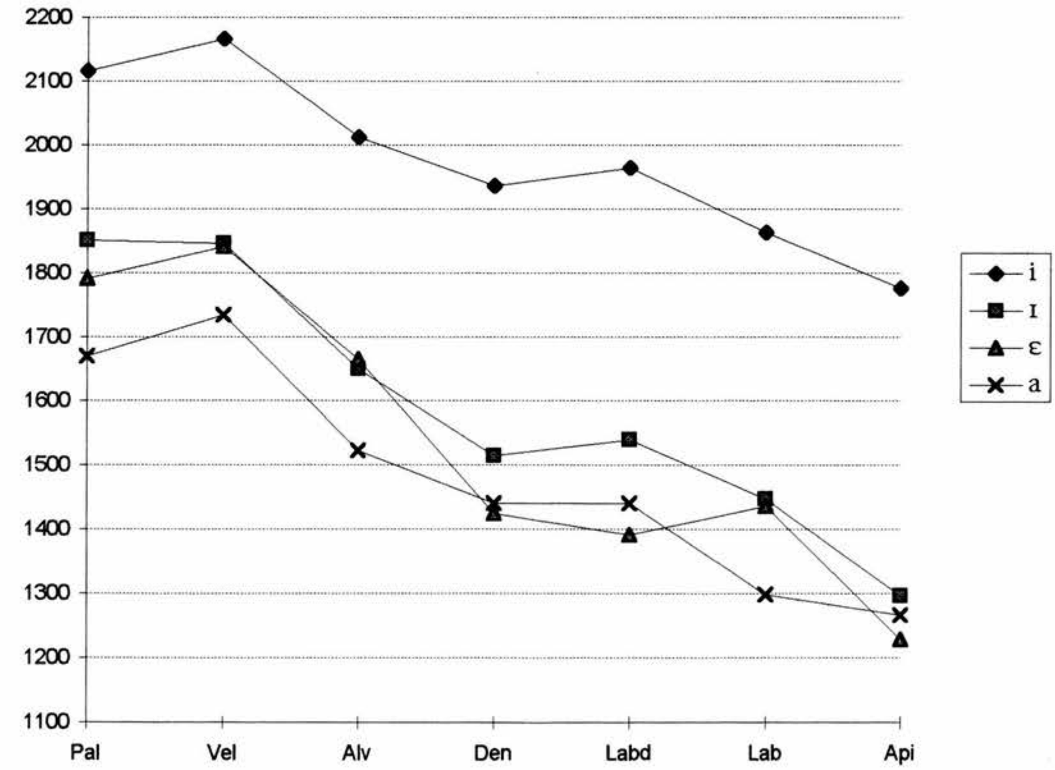
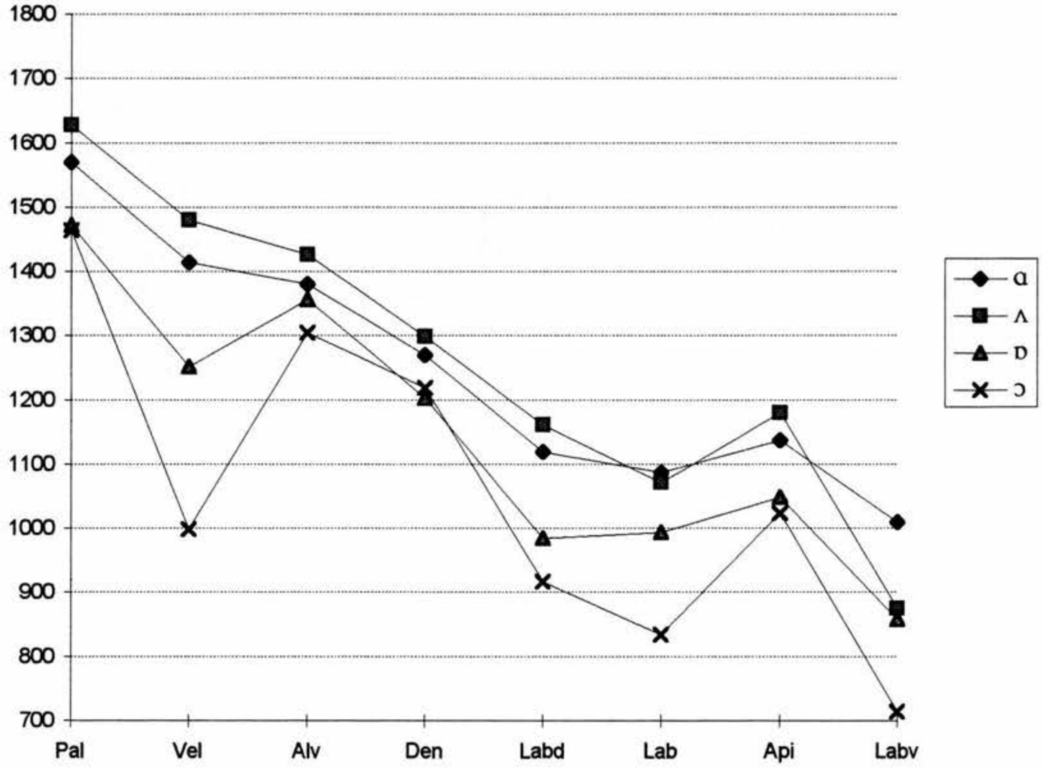


Figure 5.37: F2 mean onset/offset values as a function of C^1/C^2 place of articulation - back vowels

5.37a: mean onset values



5.37b: mean offset values

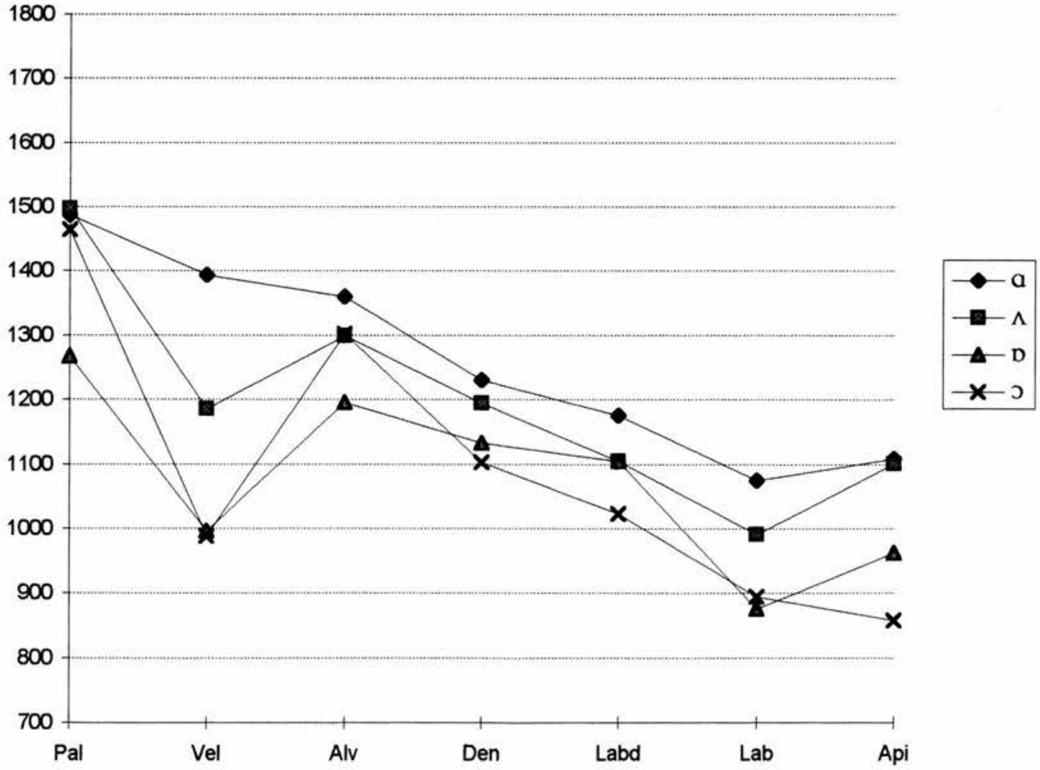
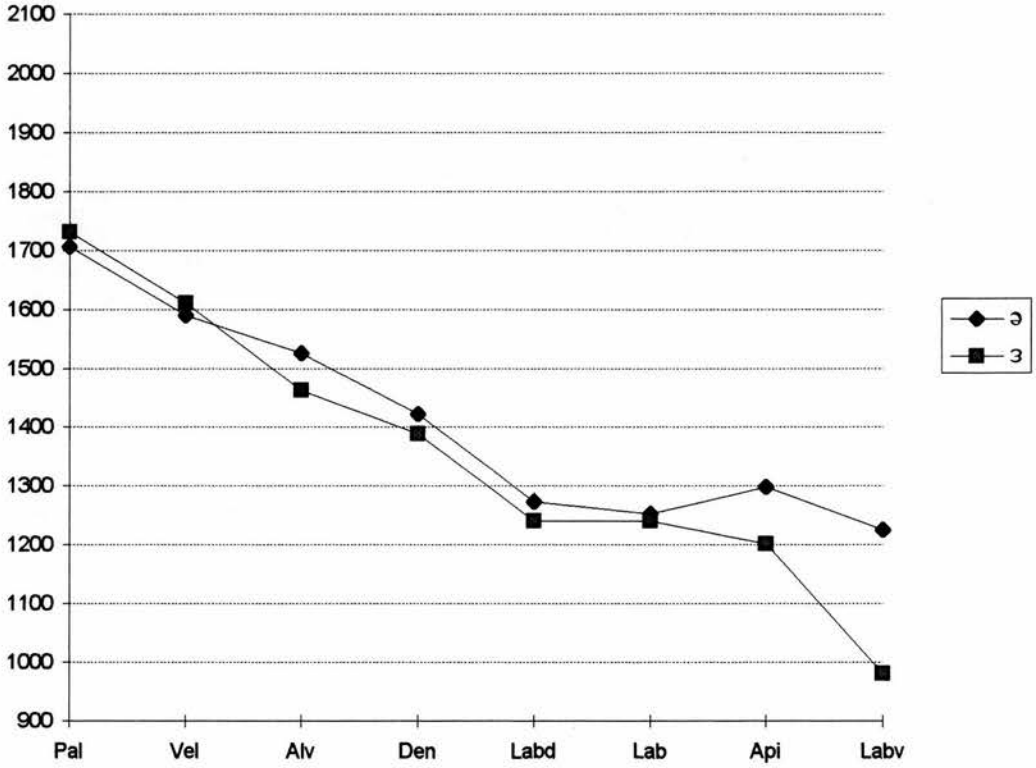


Figure 5.38: F2 mean onset/offset values as a function of C^1/C^2 place of articulation - /ə/ & /ɜ/

5.38a: mean onset values



5.38b: mean offset values

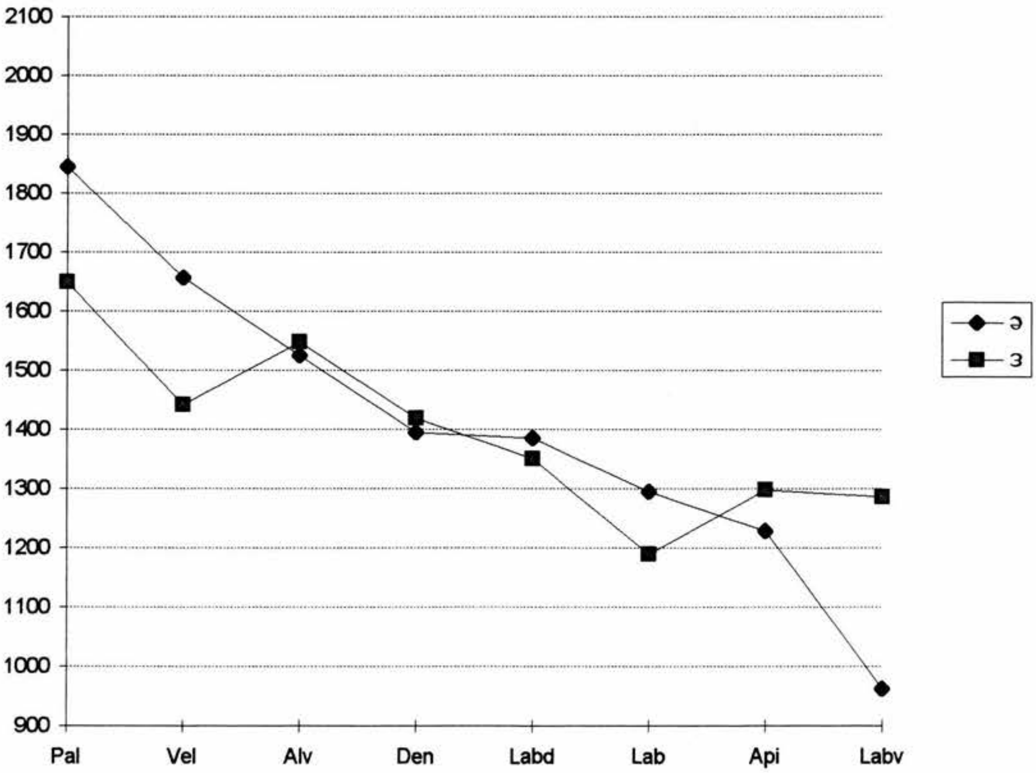
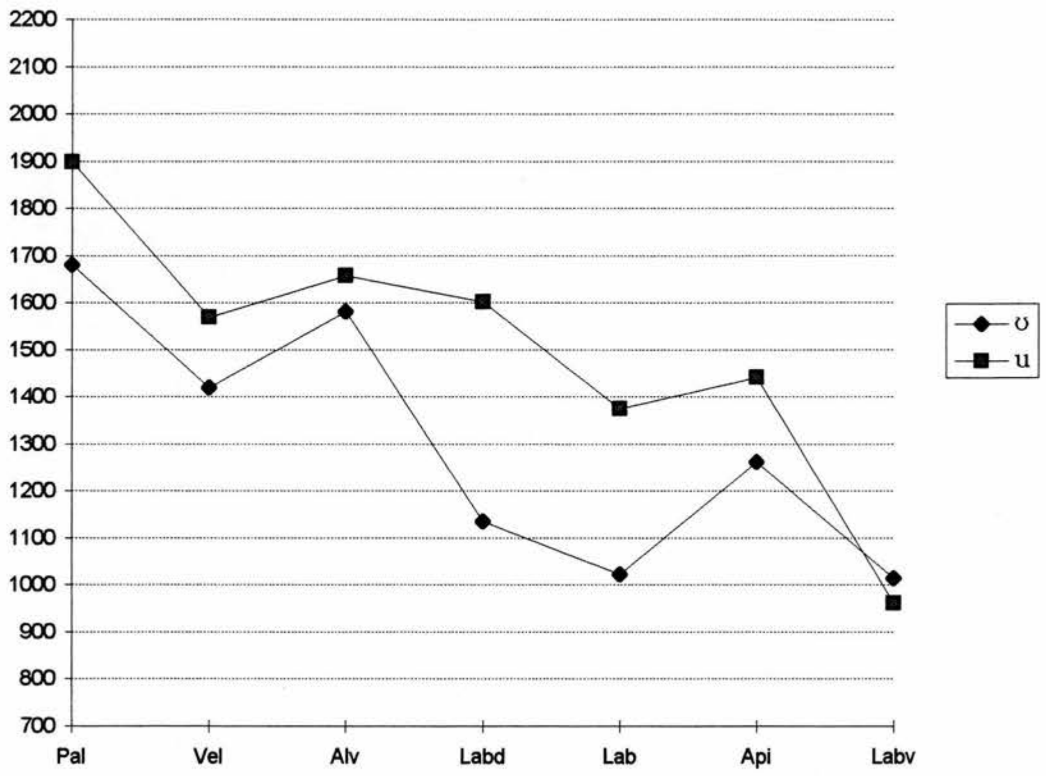


Figure 5.39: F2 mean onset/offset values as a function of C^1/C^2 place of articulation - /ʊ/ & /u/

5.39a: mean onset values



5.39b: mean offset values

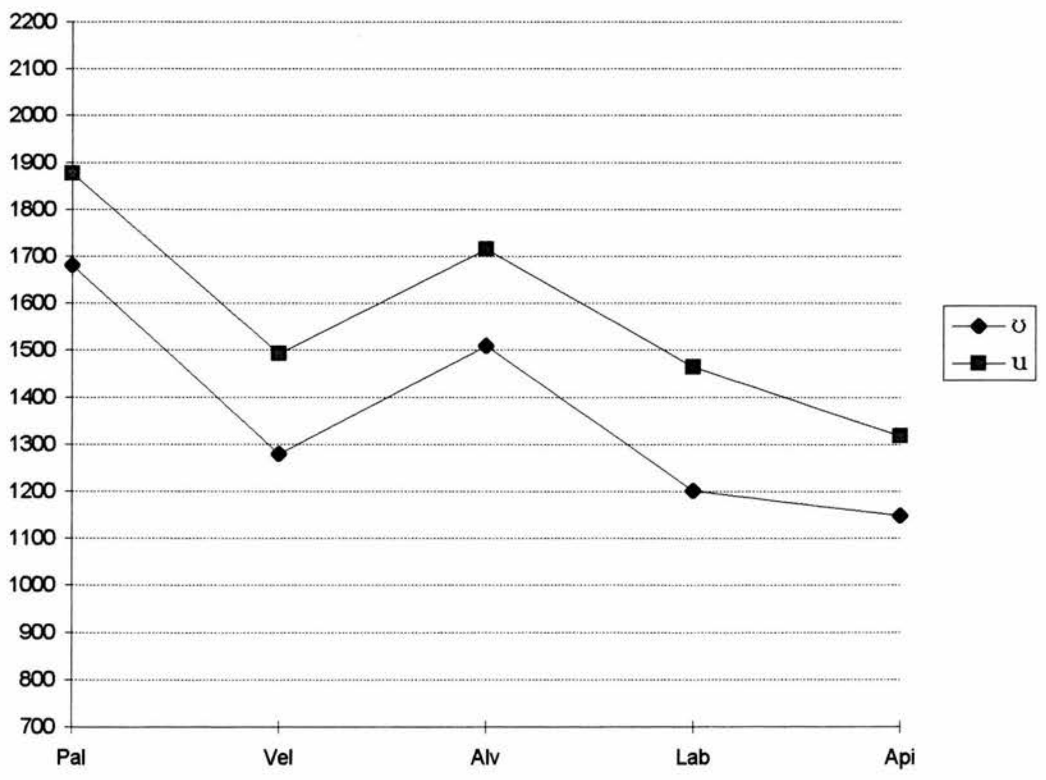
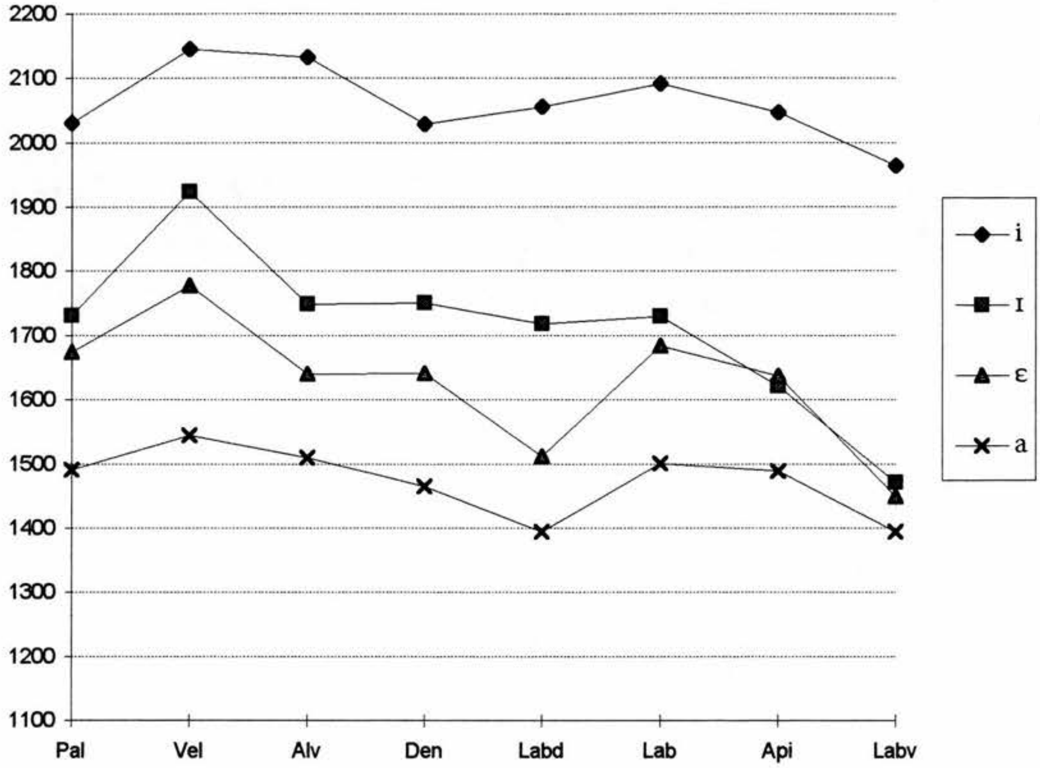


Figure 5.40: F2 mean midpoint values as a function of C^1/C^2 place of articulation - front vowels

5.40a: C^1



5.40b: C^2

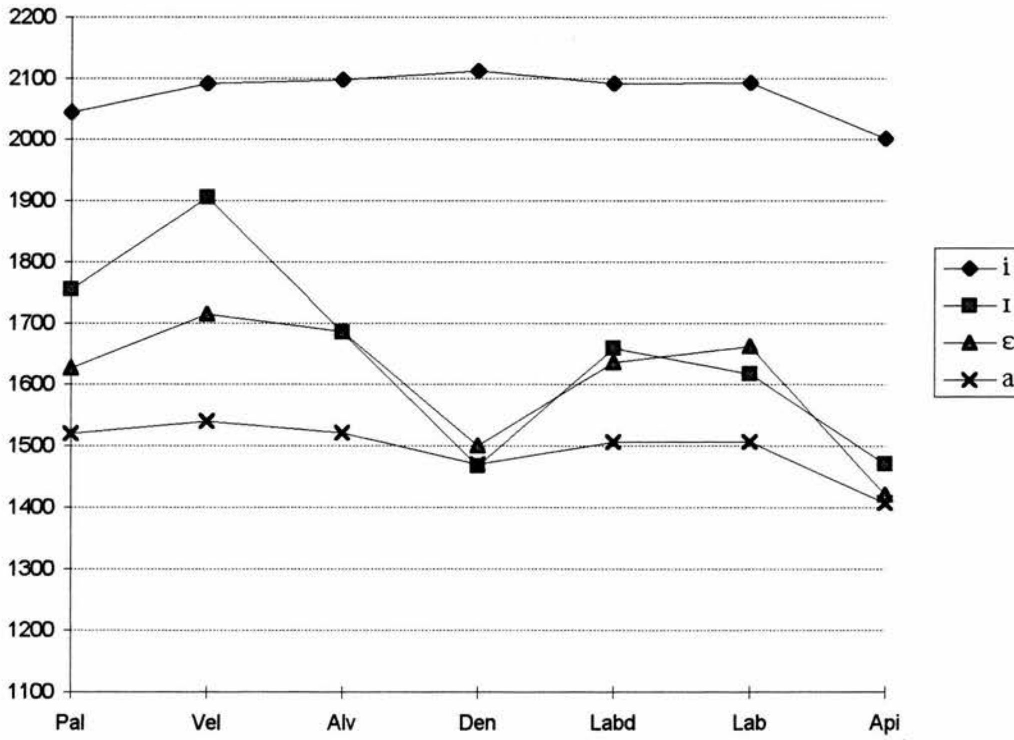
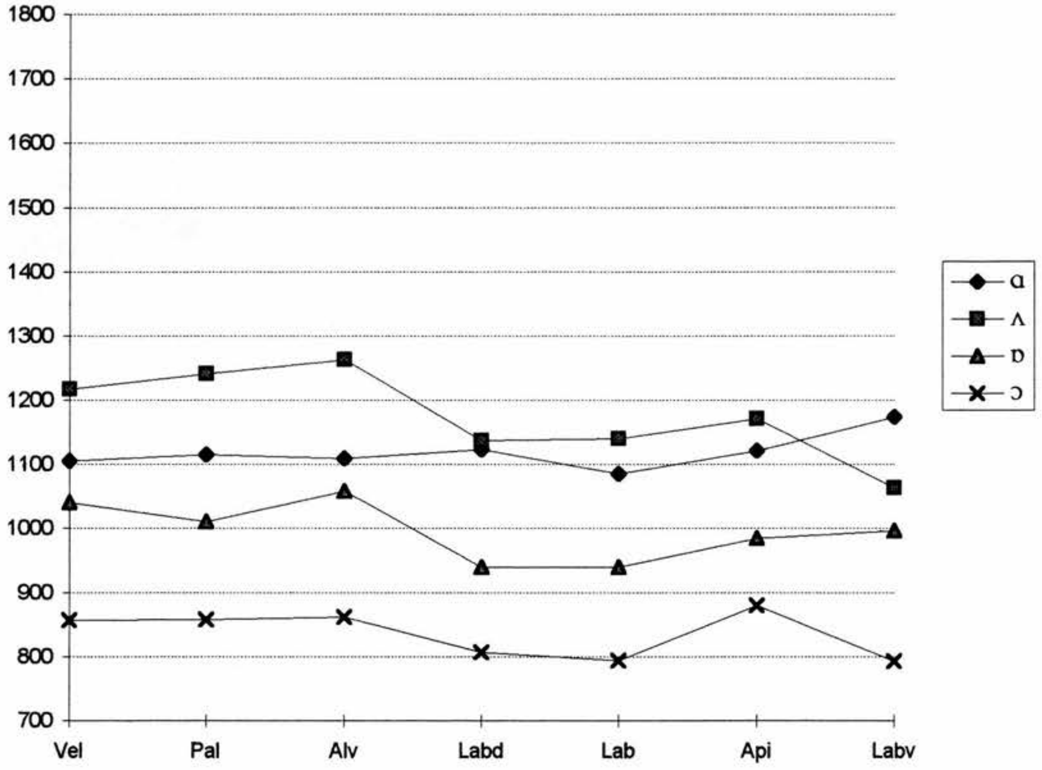


Figure 5.41: F2 mean midpoint values as a function of C^1/C^2 place of articulation - back vowels

5.41a: C^1



5.41b: C^2

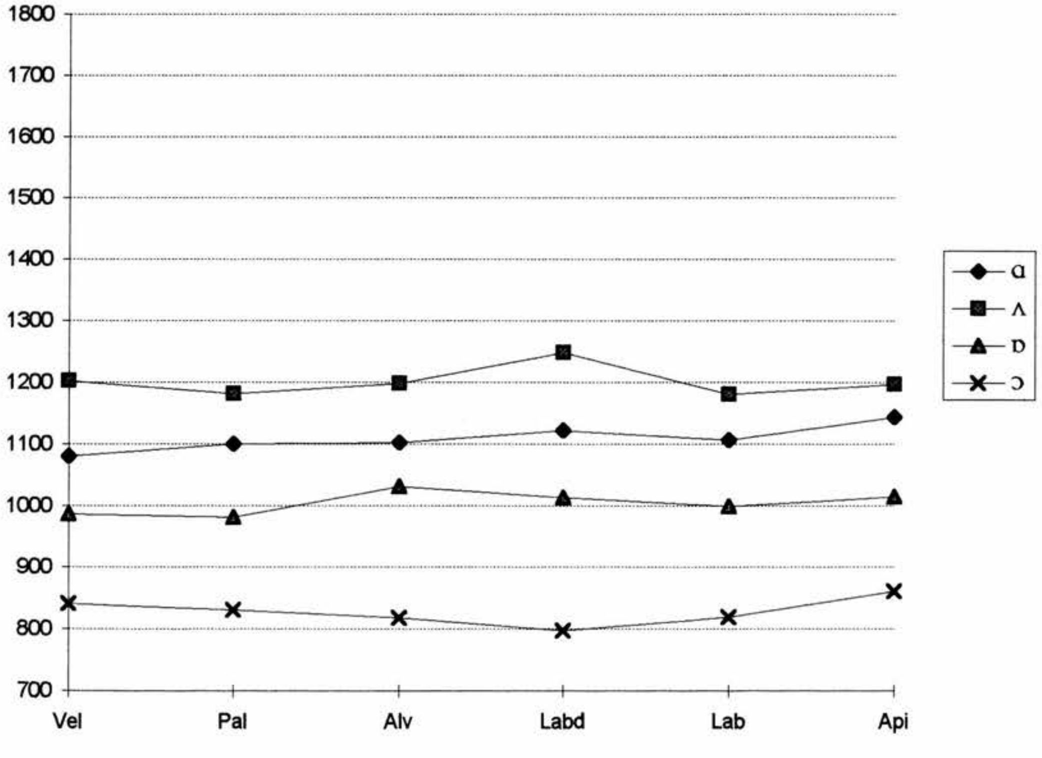
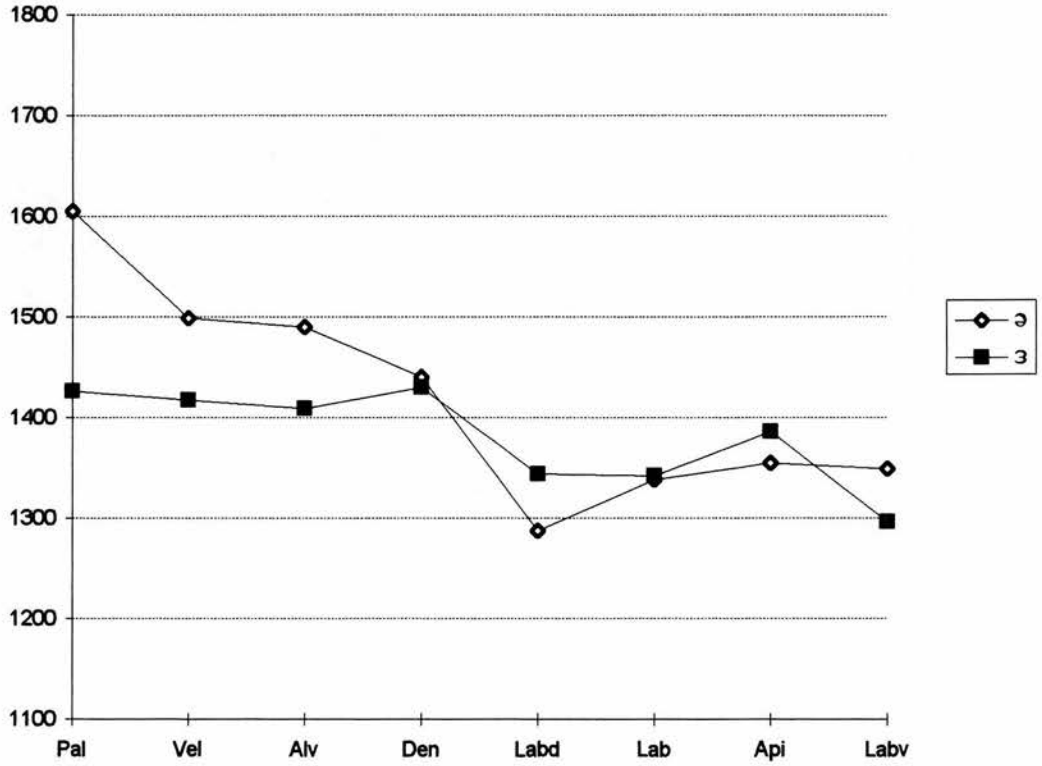


Figure 5.42: F2 mean midpoint values as a function of C^1/C^2 place of articulation - /ə/ & /ɜ/

5.42a: C^1



5.42b: C^2

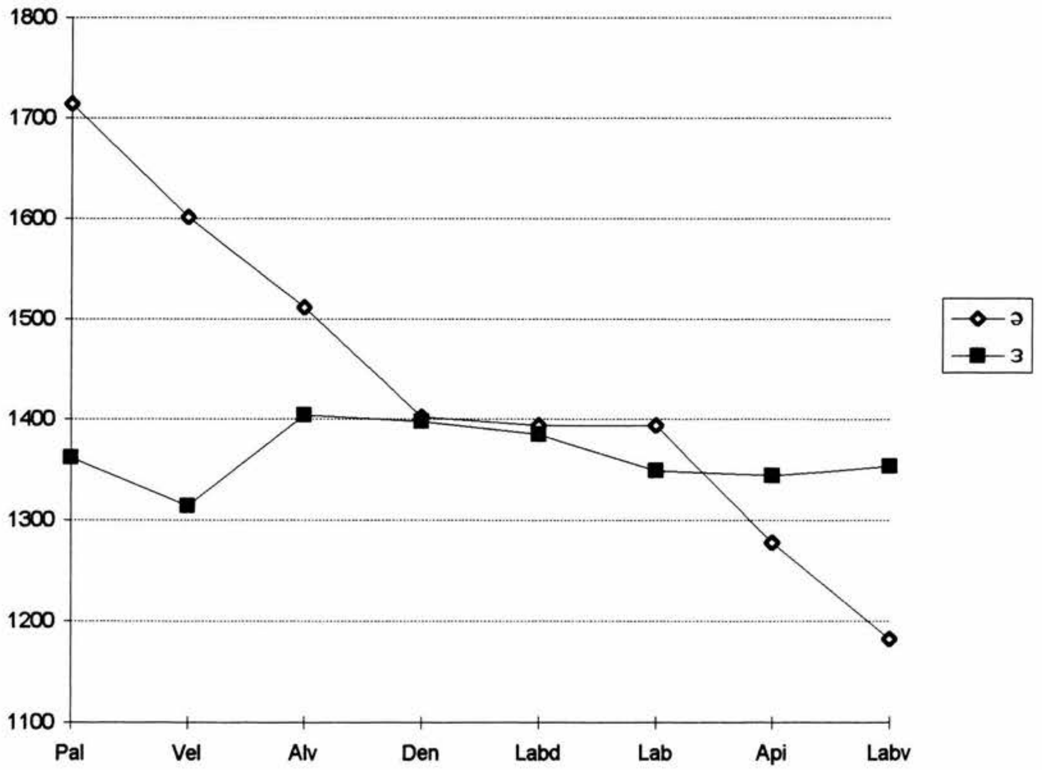
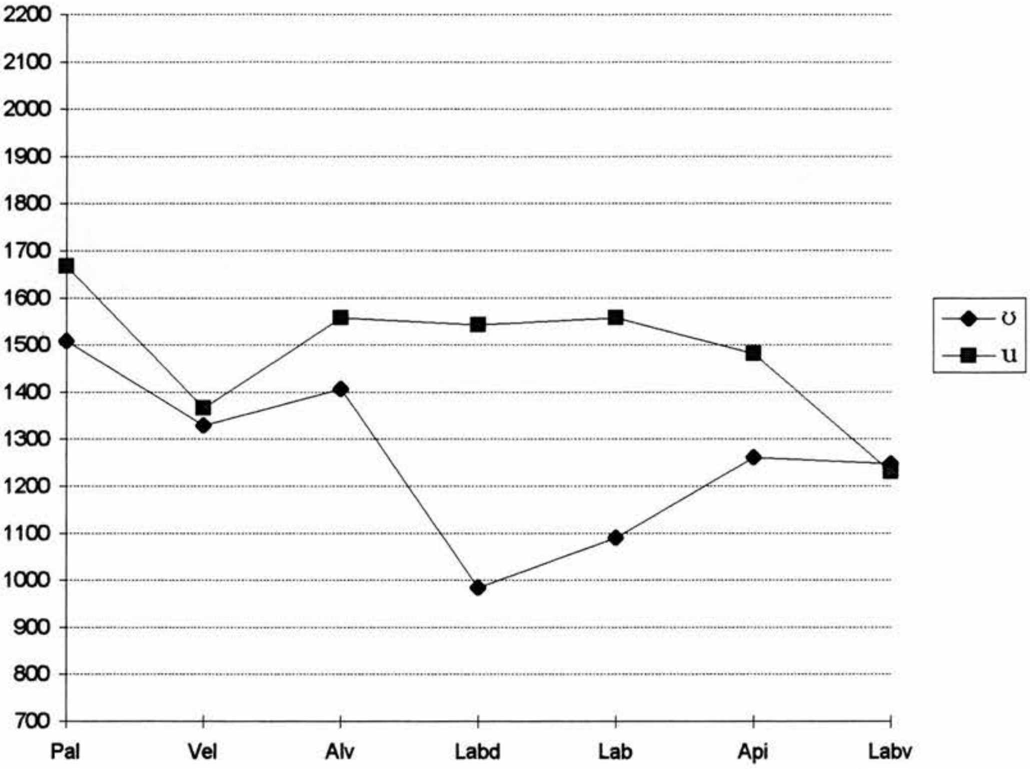


Figure 5.43: F2 mean midpoint values as a function of C^1/C^2 place of articulation - /ʊ/ & /u/

5.43a: C^1



5.43b: C^2

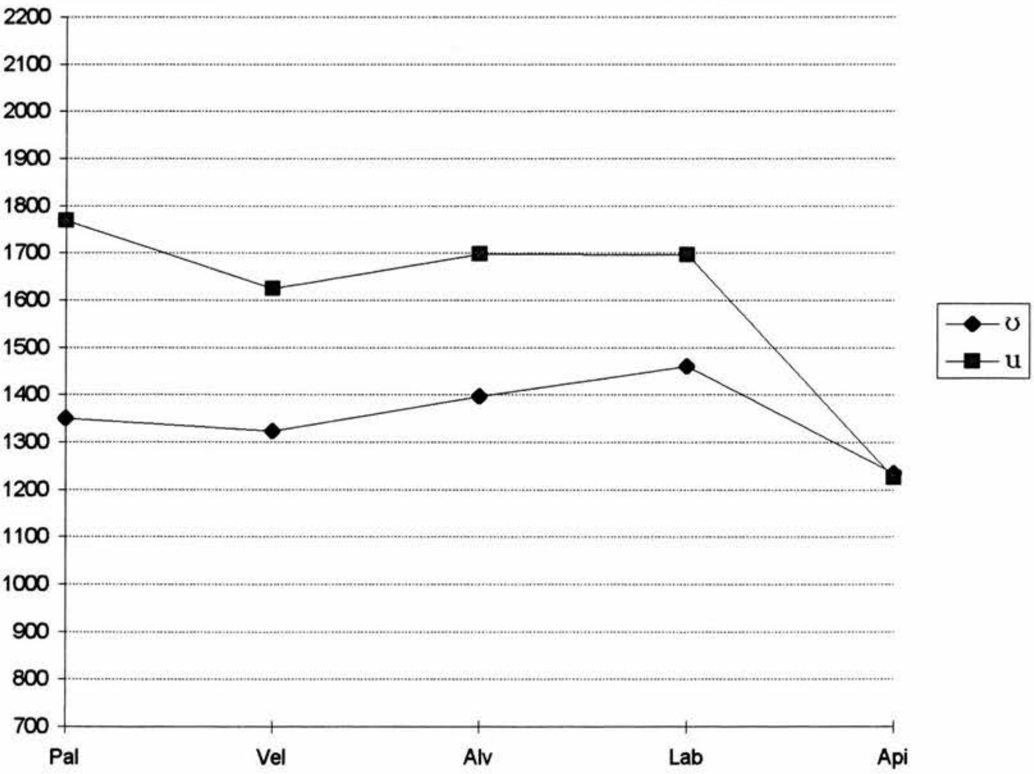
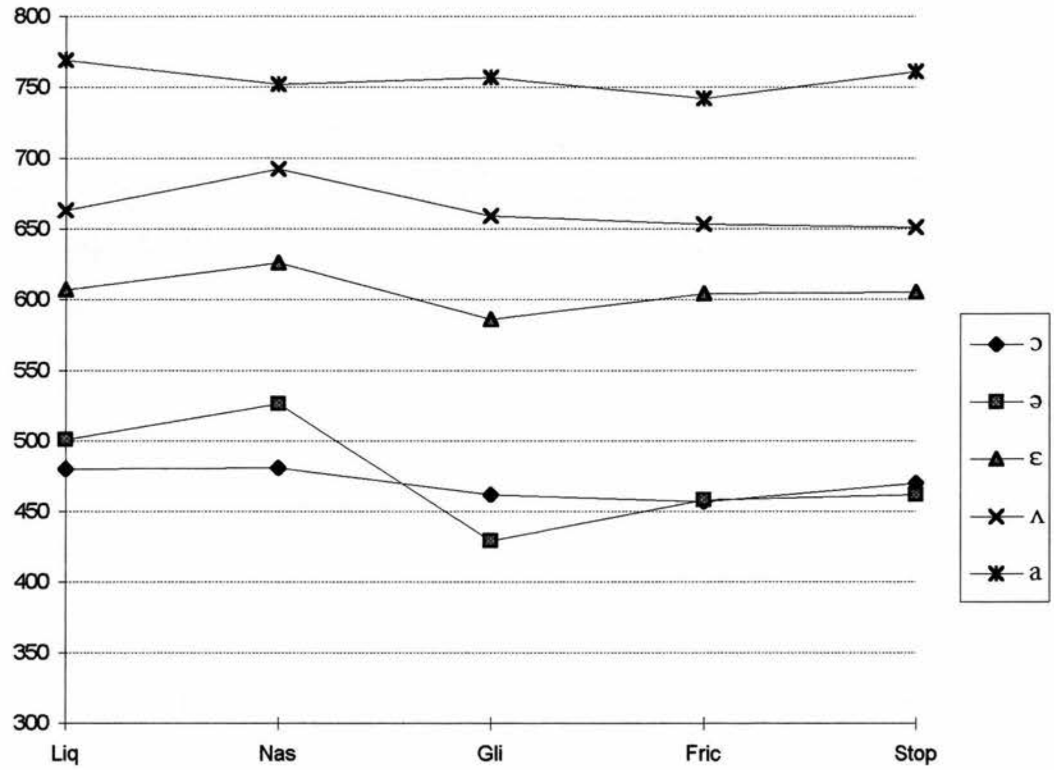


Figure 5.44: F1 mean midpoint values as a function of C^1/C^2 manner of articulation

5.44a: C^1



5.44b: C^2

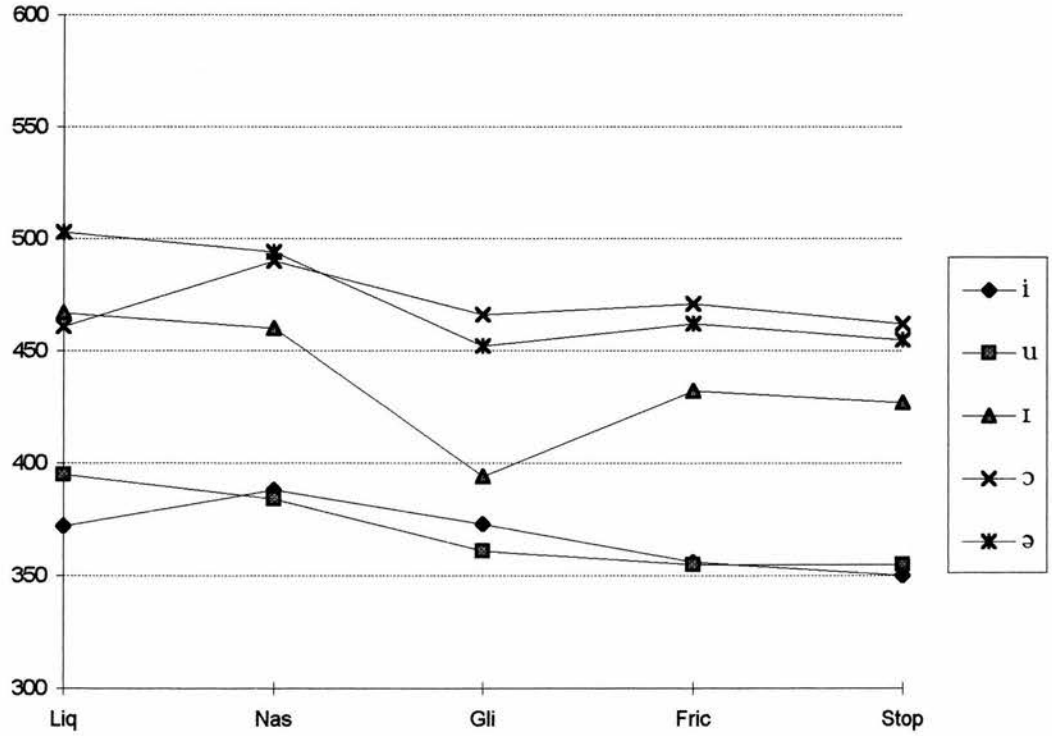


Figure 5.45: Mean second formant trajectories for /ɑ/

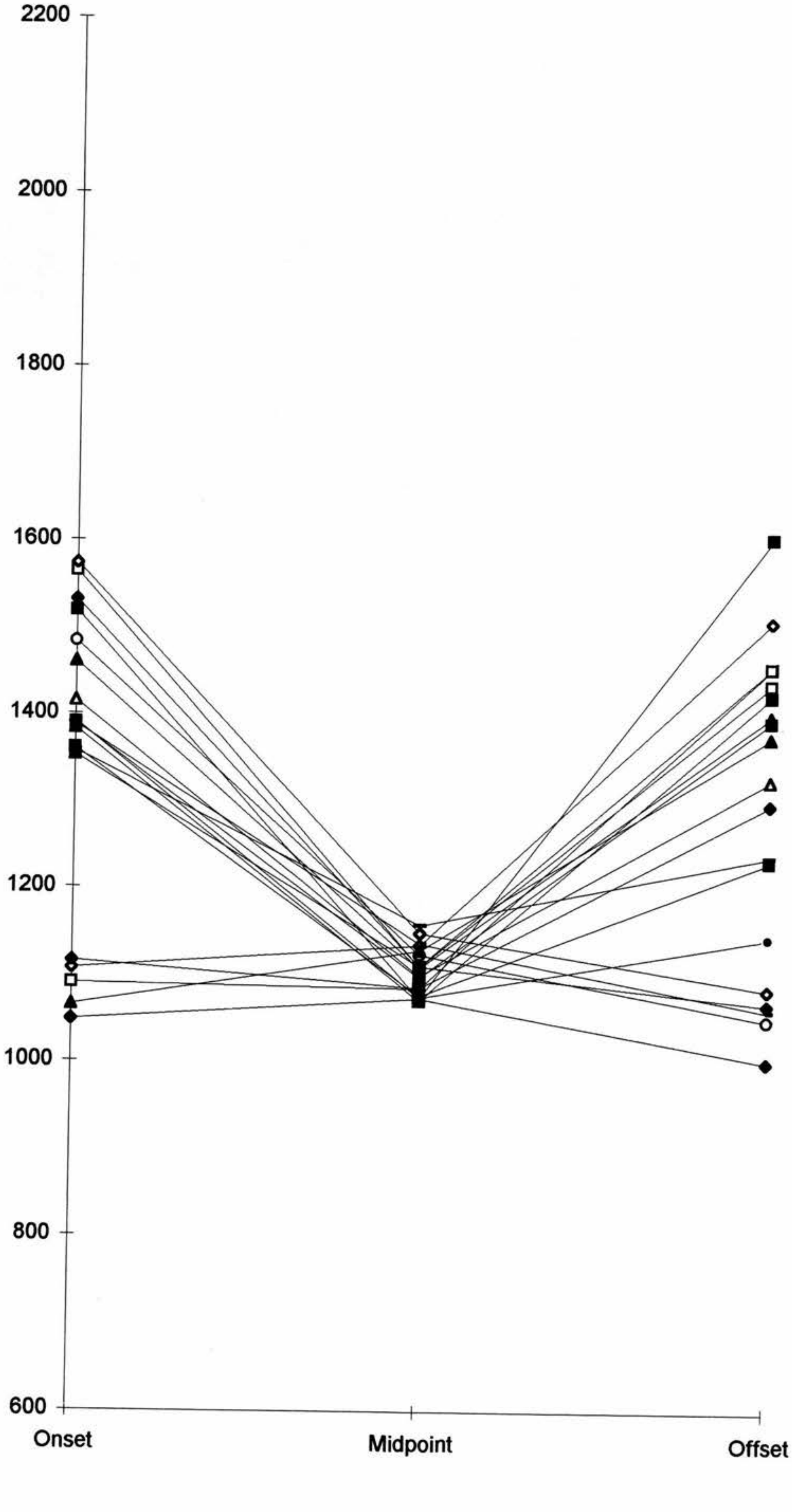


Table 5.1: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ə/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	1810	1786	1788	17	1713	1633	1579	17	*	*	*	*
VEL	1817	1779	1795	36	*	*	*	*	1693	1625	1539	12	*	*	*	*
ALV	1706	1762	1870	36	1618	1678	1736	1	1546	1529	1519	5	1498	1458	1409	6
DEN	1591	1734	1839	25	1503	1599	1655	27	1457	1489	1518	2	1342	1374	1451	30
LABD	*	*	*	*	1304	1376	1450	1	1368	1429	1497	5	1287	1328	1334	23
LAB	1304	1465	1830	136	1348	1500	1596	37	1302	1429	1495	41	1300	1383	1436	20
API	1370	1545	1774	36	1251	1337	1413	7	1357	1467	1554	15	*	*	*	*
LABV	1593	1685	1838	41	*	*	*	*	1244	1373	1449	35	*	*	*	*

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	1715	1657	1619	13	1772	1621	1375	63	1622	1494	1391	17	1689	1445	1047	103
VEL	1519	1409	1294	3	1672	1501	1507	118	1291	1173	1130	50	*	*	*	*
ALV	1508	1449	1402	8	1494	1439	1346	25	1449	1350	1262	7	1440	1236	965	45
DEN	1408	1400	1450	39	1390	1376	1312	33	1344	1345	1337	6	1257	1163	983	57
LABD	1232	1269	1307	1	1269	1214	1166	5	1185	1145	1143	25	1028	958	781	71
LAB	1260	1268	1371	63	1227	1205	1122	41	1066	1077	1086	1	1239	1160	1050	21
API	1252	1258	1287	15	1240	1234	1159	46	1218	1219	1195	17	1221	1160	946	102
LABV	943	1083	1263	27	929	982	989	27	995	1065	1134	1	*	*	*	*

Table 5.2: Coefficient of variance values for F1 and F2 at vowel onset, midpoint and offset.

	F1			*F2*		
Vowel	*Onset*	*Midpoint*	*Offset*	*Onset*	*Midpoint*	*Offset*
i	12.82	11.31	22.41	9.65	4.87	10.7
a	16.09	6.22	17.92	15.07	5.85	15.15
ɜ	13.3	6.52	17.14	16.98	5.6	11.47
ɑ	13.86	7.6	21.33	17.25	6.43	16.67
ɔ	15.98	7.23	25.47	27.22	10.06	25.55
u	11.3	9.71	16.62	16.55	15.7	19.37
ɪ	15.2	13.94	18.07	14.45	12.81	15.25
ɛ	14.04	8.86	19.88	15.75	9.68	15.61
ə	16.44	15.17	19.48	14.97	13.42	15.71
ʌ	15.49	9.63	20.18	18.76	9.16	15.53
ɒ	14.47	10.61	24.07	20.89	11.08	19.41
ʊ	18.8	8.99	12.75	23.31	21.08	19.55

Table 5.3: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ə/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	404	408	434	15	*	*	*	*	445	414	381	1	475	449	405	12	415	423	447	11
Pal.gli	*	*	*	*	*	*	*	*	367	349	319	8	*	*	*	*	423	436	426	15	369	396	386	25	371	395	465	31
Vel.stop	454	449	475	21	*	*	*	*	*	*	*	*	590	515	437	2	490	448	410	3	501	474	435	8	494	432	431	34
Vel.nas	*	*	*	*	*	*	*	*	395	482	442	85	*	*	*	*	429	575	493	152	505	586	592	50	570	521	477	3
Alv.stop	492	464	526	60	370	352	332	1	449	424	395	3	*	*	*	*	469	434	405	4	500	480	430	20	463	452	452	7
Alv.nas	361	466	433	92	*	*	*	*	*	*	*	*	*	*	*	*	473	508	478	43	485	571	483	116	435	482	484	30
Alv.fric	414	433	459	5	*	*	*	*	*	*	*	*	*	*	*	*	421	416	394	11	453	452	417	23	456	453	449	1
Den	440	453	517	34	409	384	351	5	461	449	422	10	*	*	*	*	442	444	423	15	481	499	454	42	437	437	466	19
Labden	*	*	*	*	432	401	375	3	479	451	400	15	*	*	*	*	468	454	417	15	513	504	447	32	497	510	526	2
Lab.stop	547	574	697	64	*	*	*	*	400	404	342	44	*	*	*	*	474	475	442	23	495	494	450	29	479	456	422	7
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	453	487	423	65	460	513	454	75	479	526	508	43
Api	459	443	480	35	528	466	406	1	484	482	432	32	*	*	*	*	488	468	465	11	552	565	495	55	494	505	547	21
Labvel	405	393	418	25	363	349	351	11	409	382	438	55	*	*	*	*	418	410	375	18	461	467	433	27	*	*	*	*

Table 5.3: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ə/ (continued).

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	395	411	400	18	389	370	365	9	435	442	393	37	491	546	530	47	461	459	469	8	453	466	436	29
Pal.gli	*	*	*	*	367	369	370	1	411	414	396	14	445	455	439	17	452	468	464	13	363	364	357	5
Vel.stop	*	*	*	*	463	443	425	1	472	493	418	64	576	500	426	1	491	479	474	5	511	529	489	39
Vel.nas	583	659	608	85	*	*	*	*	*	*	*	*	*	*	*	*	396	575	607	98	*	*	*	*
Alv.stop	465	434	411	5	492	467	463	14	463	449	418	11	503	486	450	13	509	473	455	12	507	444	394	9
Alv.nas	464	591	430	192	430	517	502	68	463	516	424	97	480	505	436	63	502	580	540	79	475	655	402	289
Alv.fric	*	*	*	*	441	441	444	2	466	442	411	5	423	414	399	4	456	465	461	9	474	464	429	17
Den	446	465	526	28	465	468	549	52	456	457	416	28	485	485	427	39	502	510	481	25	466	448	398	21
Labden	469	451	438	3	485	523	588	18	490	472	392	41	514	480	415	21	522	500	486	5	466	462	403	37
Lab.stop	523	500	553	51	490	491	589	65	479	465	406	30	536	494	381	47	490	480	465	3	607	549	499	5
Lab.nas	*	*	*	*	523	574	589	24	513	482	368	55	*	*	*	*	448	492	491	30	*	*	*	*
Api	419	424	474	30	487	490	503	7	524	518	443	46	486	489	443	33	526	536	498	32	509	509	450	39
Labvel	384	386	403	10	395	454	490	15	420	432	399	30	418	418	418	0	425	440	438	11	506	523	459	54

Table 5.4: Separate set analysis: Regression results for F1

	β value								R^2	*F-value*	*df.*
	V^1 *mid*[1]	V^1 *off*[2]	C^1	C^2	V^2	*Dur*	V^1 *dur*	V^2 *dur*			
CC əCC			.43**	.51**		.16**			**.7616**	**426.05**	400
CC əCV			.62**	.41**	.04*	.07*		.03	**.8094**	**578.26**	681
VC əCC	.08**		.54**	.34**		-.05	.18**		**.6412**	**161.53**	452
VC əCV	.13*		.49**	.37**	.09**	.01	.09**	.20**	**.6951**	**188.86**	580
CV əCC^{1}	.36**		-.08	.67**		.14**	.05		**.7887**	**58.21**	78
CV əCC^{2}		.55**	-.09	.50**		.09*	.06		**.8767**	**110.90**	78
CV əCV^{1}	.35**		.14	.40**	.11	.33**	-.11	-.05	**.6301**	**17.81**	71
CV əCV^{2}		.65**	-.05	.37**	.08	.29**	-.08	.00	**.7843**	**36.87**	71

[1] V^1 measured at the vowel midpoint
[2] V^1 measured at the vowel offset
** denotes significance at $p < 0.01$ and * denotes significance at $p < .05$

Table 5.5: Separate set analysis: Regression results for F2

	β value								R^2	*F-value*	*df.*
	V^1 *mid*[1]	V^1 *off*[2]	C^1	C^2	V^2	*Dur*	V^1 *dur*	V^2 *dur*			
CC əCC			.60**	.48**		.04**			**.9039**	**1254.78**	400
CC əCV			.54**	.54**	.02*	.00		.01	**.9350**	**1958.75**	681
VC əCC	.02		.47**	.59**		.05**	.01		**.9011**	**823.86**	452
VC əCV	.04*		.57**	.54**	-.01	.00	-.01	.02	**.9107**	**844.96**	580
CV əCC^{1}	.50**		.29*	.69**		-.09	.06		**.6928**	**35.18**	78
CV əCC^{2}		.65**	.10*	.45**		.01	-.01		**.8713**	**105.63**	78
CV əCV^{1}	.36**		-.20	.74**	-.06	-.07	.04	-.07	**.7328**	**27.81**	71
CV əCV^{2}		.69**	.23**	.48**	-.02	-.05	-.09	-.01	**.9182**	**113.80**	71

Table 5.6: Regression results for F2

'lhs' denotes preceding context, 'rhs', following context and 'dur', duration.
β values which fail to attain significance at $p < .01$ are italicized and emboldened

	R^2	*F-value*	*df.*		*β value*	
				lhs	*rhs*	*dur*
i	**.5294**	**246.77**	658	.48	.32	.44
ɪ	**.8151**	**1980.19**	1348	.62	.41	.17
ɛ	**.6973**	**343.94**	448	.52	.53	.18
a	**.3685**	**63.22**	325	.48	.40	.11
ə	**.9166**	**8057.98**	2199	.57	.52	.02
ɜ	**.5192**	**51.48**	143	.61	.36	***.02***
ɑ	**.1119**	**6.69**	181	.24	***-.10***	-.21
ʌ	**.5785**	**129.94**	284	.71	.40	-.18
ɔ	**.3770**	**50.22**	249	.49	***.05***	-.46
ɒ	**.4449**	**91.63**	343	.56	.18	-.43
ʊ	**.9024**	**283.53**	92	.57	.49	-.16
u	**.7569**	**210.64**	203	.52	.53	***.04***

Table 5.7: Regression results for F1

	R^2	*F-value*	*df.*		*β value*	
				lhs	*rhs*	*dur*
i	**.3923**	**141.61**	658	.45	.36	-.17
ɪ	**.6352**	**778.97**	1342	.50	.46	.07
ɛ	**.4209**	**109.02**	450	.35	.46	.38
a	**.3633**	**61.44**	323	.50	.22	.54
ə	**.7372**	**2049.44**	2192	.55	.43	.17
ɜ	**.3818**	**29.43**	143	.29	.31	.44
ɑ	**.1320**	**8.77**	173	.31	.15	.23
ʌ	**.4355**	**72.77**	283	.39	.33	.49
ɔ	**.0650**	**5.76**	249	.23	***-.11***	***-.02***
ɒ	**.3359**	**57.66**	342	.41	.32	.32
ʊ	**.3428**	**16**	92	***.13***	.47	.22
u	**.4514**	**55.69**	203	.28	.52	***-.14***

Table 5.8: Pooled data analysis excluding extreme outliers: Regression results for F2. 'lhs' denotes preceding context, 'rhs', following context and 'dur', duration. β values which fail to attain significance at p < .01 are italicized and emboldened.

	R^2	F-value	df.		β value	
				lhs	rhs	dur
i	**.5817**	**296.2**	639	.51	.35	.44
ɪ	**.8963**	**3747.83**	1301	.57	.50	.14
ɛ	**.7284**	**394.24**	441	.52	.57	.16
a	**.3685**	**63.22**	325	.48	.40	.11
ə	**.9463**	**12620.96**	2121	.43	.51	.16
ɜ	**.5240**	**51.75**	141	.61	.36	***.04***
ɑ	**.1243**	**8.14**	172	.24	-.19	-.19
ʌ	**.6179**	**146.6**	272	.76	.44	-.11
ɔ	**.3555**	**43.94**	239	.52	***.05***	-.46
ɒ	**.5079**	**113.86**	331	.61	.18	-.47
ʊ	**.9024**	**283.53**	92	.57	.49	-.16
u	**.7740**	**228.35**	200	.55	.50	.07

Table 5.9: Pooled data analysis excluding extreme outliers: Regression results for F1.

	R^2	F-value	df.		β value	
				lhs	rhs	dur
i	**.5564**	**265.92**	636	.50	.45	-.17
ɪ	**.7328**	**1194.72**	1307	.53	.48	.08
ɛ	**.4855**	**139.37**	443	.40	.46	.42
a	**.3608**	**59.27**	315	.49	.21	.55
ə	**.8088**	**2984.25**	2116	.57	.44	.15
ɜ	**.3689**	**27.86**	143	.27	.31	.44
ɑ	**.1320**	**8.76**	173	.31	.15	.23
ʌ	**.3943**	**60.75**	280	.36	.32	.49
ɔ	**.0773**	**6.64**	238	.24	***.12***	***.02***
ɒ	**.3510**	**60.02**	333	.42	.32	.35
ʊ	**.3428**	**16**	92	***.13***	.47	.22
u	**.5907**	**95.73**	199	.37	.56	-.10

Table 5.12: One-way Anova results for F2, C^1/C^2 place of articulation.
Results are given for F2 at vowel onset, midpoint and offset.

		Onset			*Midpoint*			*Midpoint*			*Offset*	
	C^1						*C^2*					
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	46.06	7, 248	.0000	24.38	7, 637	.0000	5.55	7, 481	.0000	33.8	7, 111	.0000
ɪ	144.25	7, 564	.0000	45.37	7, 702	.0000	74.24	7, 39	.0000	100.72	7, 31	.0000
ɛ	177.21	7, 182	.0000	21.31	7, 84	.0000	49.72	6, 39	.0000	149.02	6, 67	.0000
a	104.22	7, 48	.0000	8.08	7, 290	.0000	16.78	6, 88	.0000	90.05	6, 87	.0000
ə	192.59	7, 873	.0000	62.66	7, 1031	.0000	120.25	7, 500	.0000	314.84	7, 620	.0000
ɜ	95.39	7, 132	.0000	20.28	7, 85	.0000	4.15	7, 133	.0004	33.32	7, 133	.0000
ɑ	58.27	7, 78	.0000	1.05	7, 155	**.3990**	1.08	7, 172	**.3771**	12.5	7, 172	.0000
ʌ	116.69	7, 113	.0000	17.21	7, 271	.0000	1.19	6, 281	**.3130**	44.09	6, 281	.0000
ɒ	102.56	7, 136	.0000	10.85	7, 151	.0000	2.04	6, 134	**.0650**	51.7	6, 68	.0000
ɔ	86.73	7, 121	.0000	9.35	7, 219	.0000	2.11	7, 226	**.4836**	54.83	7, 25	.0000
ʊ	25.73	6, 39	.0000	10.31	6, 41	.0000	0.9	5, 76	**.4836**	8.82	5, 76	.0000
u	30.95	7, 17	.0000	6.19	7, 14	.0019	10.55	7, 47	.0000	22.17	7, 17	.0000

Table 5.13: Significant pairwise differences ($p < .05$) in vowel F2 onset/offset value as a function of C^1/C^2 place of articulation.

Significant differences are indicated by inclusion of the appropriate vowel symbol within a given cell.

5.13a: C^1

	Pal	*Vel*	*Alv*	*Den*	*Labd*	*Lab*	*Api*	*Labv*
Pal	*	ə i ɪ ɛ a ɑ ɒ ɔ ʊ	ə i ɪ ɛ a ɑ ʌ ɒ ɜ u	ə i ɪ a ɑ ʌ ɒ ɜ u	ə i ɪ ɛ a ɑ ʌ ɒ ɔ ʊ ɜ	ə i ɪ ɛ a ɑ ʌ ɒ ɔ ʊ ɜ u	ə i ɪ ɛ a ɑ ʌ ɒ ɔ ʊ ɜ u	ə i ɪ ɛ ʌ ɒ ɔ ʊ ɜ u
Vel		*	i ɪ ɛ a ɑ ɔ	ə i ɪ ɛ a ɑ ʌ ɜ	ə i ɪ ɛ a ɑ ʌ ɒ ɜ	ə i ɪ ɛ a ɑ ʌ ɒ ɔ ʊ ɜ	ə i ɪ ɛ a ɑ ʌ ɒ ɜ	ə i ɪ ɛ ʌ ɒ ɔ ɜ
Alv			*	ə i ɪ a ɑ ɒ	ə i ɪ ɛ a ɑ ʌ ɒ ɔ ɜ	ə i ɪ ɛ a ɑ ʌ ɒ ɔ ʊ ɜ u	ə i ɪ ɛ a ɑ ʌ ɒ ɔ ʊ ɜ u	ə i ɪ ɛ ʌ ɒ ɔ ʊ ɜ
Den				*	ə ɒ ɜ	ə ɑ ʌ ɒ ɜ u	ə ɪ ɑ ɒ ɜ u	ə i ɪ ɛ ʌ ɒ ɜ
Labd					*	ɛ	ɪ ʊ	ɪ ɛ ʌ ɒ ɔ ɜ
Lab						*	i ɛ ɔ	i ɪ ɛ ɒ ɔ ɜ
Api							*	ə i ɛ ɒ ɔ ɜ
Labv								*

5.13b: C^2

	Pal	*Vel*	*Alv*	*Den*	*Labd*	*Lab*	*Api*	*Labv*
Pal	*	ə ʌ ʊ u	ə ɪ ɛ a	ə ɪ ɛ ʌ u	ə ɪ ɛ ɑ ʌ ɜ u	ə i ɪ ɛ a ɑ ʌ ɒ ɜ ʊ u	ə i ɪ ɛ a ɑ ɜ ʊ u	ə i ɑ
Vel		*	ə i ɪ ɛ a ʌ ɒ ɔ ʊ u	ə i ɪ ɛ a ɒ	ə i ɪ ɛ a ɑ	ə i ɪ ɛ a ɑ ʌ ɒ ɜ	ə i ɪ ɛ a ɑ ɔ u	ə i ɑ ɔ ɜ
Alv			*	ə ɪ ʌ ɔ ɜ	ə ɪ ɛ ɑ ʌ ɔ ɜ u	ə i ɪ ɛ a ɑ ʌ ɒ ɔ ɜ ʊ u	ə i ɪ ɛ a ɑ ʌ ɒ ɔ ɜ ʊ u	ə i ɑ ɔ
Den				*		ə ʌ ɒ ɔ ɜ u	ə ɪ ɒ ɔ u	ə i ɔ
Labd					*	ə ɪ ʌ ɒ ɜ	ə i ɪ ɛ u	ə i ɑ
Lab						*	ə ɪ ɛ ʌ ɒ	ə i ɑ ɜ
Api							*	ə ɑ
Labv								*

Table 5.14: Significant pairwise differences ($p < .05$) in vowel F2 midpoint value as a function of C^1/C^2 place of articulation.
Significant differences are indicated by inclusion of the appropriate vowel symbol within a given cell.

5.14a: C^1

	Pal	*Vel*	*Alv*	*Den*	*Labd*	*Lab*	*Api*	*Labv*
Pal	*	ə i ɪ ɛ	ə i	ə	ə ɛ ɜ ʌ ʊ	ə ɜ ʌ ʊ	ə ɪ u ʊ	ə ɪ ɛ ɜ ʌ
Vel		*	ɪ ɛ	i ɪ a	ə i ɪ ɛ a ɜ ɒ ɔ ʊ	ə ɪ ɛ ɜ ʌ ɒ ɔ ʊ	ə i ɪ ɛ a	ə i ɪ ɛ ɜ ʌ ɔ
Alv			*	ə	ə ɛ a ɜ ʌ ɒ ɔ ʊ	ə ɜ ʌ ɒ ɔ ʊ	ə ɪ ʌ ɒ	ə ɪ ɛ ɜ ʌ ɔ
Den				*	ə	ə	ə ɪ	ə ɪ ɜ ʌ
Labd					*	ɛ a	ə ɪ ɛ ɔ ʊ	ɪ
Lab						*	ɪ ɛ a ɒ ɔ ʊ	ɪ ʌ
Api							*	ɪ ɛ ʌ ɔ
Labv								*

5.14b: C^2

	Pal	*Vel*	*Alv*	*Den*	*Labd*	*Lab*	*Api*	*Labv*
Pal	*	ɪ	ə ɪ	ə ɪ	ə ɪ	ə ɪ	ə ɪ ɛ a u	ə
Vel		*	ə ɪ ɜ ɒ	ə ɪ	ə ɪ ɛ a	ə ɪ a	ə i ɪ ɛ a u	ə
Alv			*	ə ɪ	ə ɛ	ə ɜ	ə i ɪ ɛ a u	ə
Den				*	ɪ	ɪ	ə	ə
Labd					*		ə i ɪ ɛ a u	ə
Lab						*	i ɪ ɛ a u	ə
Api							*	ə u
Labv								*

Table 5.15: Two-way Anova results for F2, C^1 place & manner of articulation
Results are given for F2 at the vowel midpoint.

	C^1 place			C^1 manner			Interaction		
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	16.26	101	.0001	17.39	101	.0000	3.44	272	.0333
ɪ	5.12	749	.0236	5.75	267	.0036	7.66	749	.0005
ɛ	0.40	195	.5256	27.27	136	.0000	0.64	195	.5284
a	6.39	69	.0014	21.04	68	.0000	3.24	162	.0421
ə	172.92	459	.0000	5.35	458	.0051	2.85	459	.0588
ɜ	14.24	13	.0023	1.38	86	.2572	0.94	86	.3944
ɑ	0.70	1	.4044	1.04	2	.3589	0.21	5	.8144
ʌ	64.48	56	.0000	1.14	56	.3284	1.73	56	.1874
ɒ	64.47	152	.0000	0.84	152	.4355	1.78	152	.1714
ɔ	19.47	23	.0002	3.36	121	.0379	0.09	121	.9139
u	1.29	62	.2599	2.33	62	.1057	0.08	62	.9232

Table 5.16: Two-way Anova results for F2, C^2 place & manner of articulation
Results are given for F2 at the vowel midpoint.

	C^2 place			C^2 manner			Interaction		
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	0.20	350	.6556	3.56	350	.0296	3.20	190	.0432
ɪ	9.52	160	.0024	0.32	160	.7245	6.60	818	.0014
ɛ	2.85	265	.0925	2.83	265	.0608	1.04	265	.3562
a	0.07	190	.7875	14.84	190	.0000	2.96	190	.0542
ə	231.17	836	.0000	9.09	837	.0001	5.29	1572	.0051
ɜ	7.36	101	.0078	0.33	101	.7210	1.56	101	.2155
ɑ	0.11	132	.7452	5.68	65	.0053	1.42	132	.2458
ʌ	0.34	196	.5601	5.56	196	.0045	2.98	196	.0529
ɒ	3.13	69	.0815	2.75	68	.0710	1.06	240	.3471
ɔ	0.00	128	.9478	1.53	128	.2208	1.79	128	.1706
u	0.62	103	.4333	0.25	103	.7793	5.27	103	.0066

Table 5.17: Two-way Anova results for F2, C^1/C^2 place of articulation.
Results are given for F2 at the vowel midpoint.

	C^1			C^2			*Interaction*		
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	27.42	2, 464	.0000	16.25	2, 464	.0000	2.1	4, 464	.0802
ɪ	33.02	2, 54	.0000	93.94	2, 1249	.0000	1.66	4, 1249	.1557
ε	63.79	2, 64	.0000	144.94	2, 412	.0000	2.31	4, 412	.0570
a	7.79	2, 291	.0056	26.8	2, 291	.0000	2.58	2, 291	.0775
ə	622.22	1, 2009	.0000	832.77	1, 818	.0000	22.02	1, 818	.0000
ɜ	25.39	2, 125	.0000	8.28	2, 125	.0004	0.85	4, 125	.4958
ɑ	0.67	2, 154	.5155	1.78	2, 154	.1715	1.01	4, 154	.4067
ʌ	36.89	2, 270	.0000	2.84	2, 270	.0602	1.73	4, 270	.1436
ɒ	9.44	2, 307	.0001	6.21	2, 65	.0034	0.35	4, 307	.8464
ɔ	9.25	2, 200	.0001	1.8	2, 17	.1951	4, 17	0.94	.4651
ʊ	10.17	2, 74	.0021	1.78	2, 74	.1586	2.12	2, 74	.1045
u	44.45	2, 147	.0000	16.35	2, 147	.0000	4.59	4, 147	.0016

Table 5.18: One-way Anova results for F1, C^1/C^2 place of articulation.
Results are given for F1 at the vowel midpoint.

		C^1			C^2	
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	4.31	7, 272	.0002	1.51	7, 482	**.1609**
ɪ	16.21	7, 611	.0000	8.71	7, 527	.0000
ɛ	4.02	7, 416	.0003	4.79	6, 446	.0001
a	3.63	7, 288	.0009	4.36	6, 319	.0003
ə	30.65	7, 1410	.0000	15.11	7, 793	.0000
ɜ	1.31	7, 132	**.2506**	0.55	7, 133	**.7979**
ɑ	1.46	7, 153	**.1842**	0.74	7, 169	**.6382**
ʌ	0.71	7, 270	**.6668**	3.29	6, 94	.0056
ɒ	1.97	7, 307	**.0590**	3.20	6, 32	.0142
ɔ	1.21	7, 219	**.3002**	1.25	7, 226	**.2752**
ʊ	3.84	6, 88	.0019	1.76	5, 76	**.1778**
u	1.1	7, 198	**.3622**	3.6	7, 148	.0013

Table 5.19: One-way Anova results for F1, C^1/C^2 manner of articulation.
Results are given for F1 at the vowel midpoint.

		C^1			C^2	
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	13.41	5, 188	.0000	18.13	5, 167	.0000
ɪ	41.29	5, 528	.0000	31.26	5, 396	.0000
ɛ	3.46	5, 448	.0044	7.43	4, 413	.0000
a	2.69	5, 321	.0211	3.47	3, 322	.0085
ə	73.4	5, 957	.0000	40.38	5, 460	.0000
ɜ	1.24	5, 141	**.2914**	1.83	5, 16	**.1644**
ɑ	.96	5, 171	**.4465**	5.13	4, 172	.0006
ʌ	4.05	5, 281	.0014	9.27	3, 237	.0000
ɒ	.61	5, 202	**.6935**	.78	3, 233	**.5069**
ɔ	3.22	5, 247	.0078	3.63	5, 247	.0034
ʊ	4.71	5, 63	.0010	1.54	4, 91	**.1971**
u	1.91	5, 201	**.0939**	14.71	5, 66	.0000

Table 5.20: Significant pairwise differences ($p < .05$) in vowel F1 midpoint value as a function of C^1/C^2 manner of articulation.

Significant differences are indicated by the inclusion of the appropriate vowel symbol within a given cell.

5.20a: C^1

	Stop	*Nas*	*Fric*	*Liq*	*Gli*
Stop	*	ə ɜ i ɪ ʌ	i ɪ	ə i ɪ	ə i
Nas	ə ɜ i ɪ ʌ	*	ə i ɪ ʌ ɔ	i	ə ɜ ɪ ɛ
Fric	i ɪ	ə i ɪ ʌ ɔ	*	ə ɪ a	ə
Liq	ə i ɪ		ə ɪ a	*	ɪ ʊ
Gli	ə i	ə ɜ ɪ ɛ	ə	ɪ ʊ	*

5.20b: C^2

	Stop	*Nas*	*Fric*	*Liq*	*Gli*
Stop	*	ə ɜ i ɪ ɛ ɑ ʌ ɔ u	ɑ	ə i ɪ ɛ	ɑ
Nas	ə ɜ i ɪ ɛ ɑ ʌ ɔ u	*	ə i ɪ u	a ʌ ɔ	ə ɪ ɜ ɑ
Fric	ɑ	ə i ɪ u	*	ə ɪ	ɑ
Liq	ə i ɪ ɛ	a ʌ ɔ	ə ɪ	*	ə ɪ ɑ
Gli	ɑ	ə ɪ ɜ ɑ	ɑ	ə ɪ ɑ	*

Table 5.21: Two-way Anova results for F1, C^1 place & manner of articulation.
Results are given for F1 at the vowel midpoint.

	C^1 place			*C^1 manner*			*Interaction*		
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	0.66	1, 109	**.4188**	18.98	2, 110	.0000	2.08	2, 272	**.1268**
ɪ	0.57	1, 236	**.4515**	58.32	2, 235	.0000	5.28	2, 745	.0053
ɛ	1.65	1, 195	**.2007**	2.71	2, 195	**.0689**	2.89	2, 195	**.0578**
a	3.26	1, 133	**.0731**	2.25	2, 133	**.1094**	0.34	2, 133	**.7152**
ə	8.38	1, 371	.0040	63.71	2, 369	.0000	13.84	2, 370	.0000
ɜ	1.38	1, 7	**.2791**	1.33	2, 86	**.2710**	2.37	2, 86	**.1000**
ɑ	0.12	1, 64	**.7254**	1.83	2, 25	**.1808**	0.93	2, 64	**.4006**
ʌ	0.02	1, 65	**.8895**	9.05	2, 124	.0002	0.75	2, 63	**.4751**
ɒ	4.85	1, 151	.0292	1.59	2, 151	**.2072**	1.24	2, 151	**.2937**
ɔ	0.04	1, 121	**.8513**	5.25	2, 121	.0065	00.74	2, 121	**.4786**
u	5.33	1, 14	.0368	16.71	2, 15	.0002	01.64	2, 12	**.2347**

Table 5.22: Two-way Anova results for F1, C^2 place & manner of articulation.
Results are given for F1 at the vowel midpoint.

	C^2 place			*C^2 manner*			*Interaction*		
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	5.05	1, 350	.0253	21.71	2, 90	.0000	1.48	2, 350	**.2292**
ɪ	9.66	1, 812	.0019	25.60	2, 199	.0000	1.51	2, 812	**.2207**
ɛ	17.34	1, 279	.0000	6.24	2, 279	.0022	2.67	2, 279	**.0713**
a	1.82	1, 190	**.1790**	0.44	2, 190	**.6466**	2.37	2, 190	**.0958**
ə	2.57	1, 1572	**.1090**	30.54	2, 828	.0000	6.53	2, 827	.0015
ɜ	0.06	1, 100	**.8112**	4.73	2, 46	.0135	0.89	2, 100	**.4159**
ɑ	1.82	1, 190	**.1790**	0.44	2, 190	**.6466**	2.37	2, 190	**.0958**
ʌ	0.86	1, 92	**.3567**	4.82	2, 90	.0103	2.24	2, 196	**.1095**
ɒ	5.44	1, 44	.0244	0.15	2, 45	**.8591**	1.39	2, 239	**.2499**
ɔ	0.00	1, 128	**.9777**	5.36	2, 128	.0058	0.75	2, 128	**.4759**
u	3.98	1, 102	.0487	11.05	2, 102	.0000	6.86	2, 102	.0016

Chapter 6

Role of Stress in Conditioning Vowel Quality

Introduction

The reduced quality of unstressed vowels relative to stressed vowels (where stress refers to either lexical stress in isolated words and/or sentence stress in connected speech) is widely documented in the literature (Tiffany, 1959; Lindblom, 1963; Kozhevnikov & Chistovich, 1965; Delattre, 1969; Verbrugge & Shankweiler,1977; Gay, 1978; Harris,1978; Tuller, Kelso & Harris, 1982; Summers, 1987; Engstrand, 1988; Fourakis, 1991). The precise nature of the reduction involved, however, remains unclear.

Traditionally, vowel reduction has been equated with articulatory and acoustic centralisation, that is, with movement toward the central or neutral vowel position associated with schwa. For example, Tiffany (1959) claims that unstressed vowels are inherently more central than stressed vowels, that is, that they are characterised by less extreme articulatory and less distinctive spectral configurations than their stressed counterparts. Tiffany compared the American English vowels /i, ɪ, e, ɛ, æ, o, ʊ, u, ɑ, ʌ, aɪ, aʊ/ produced in isolation with the corresponding vowels produced in an /h-d/ context embedded in carrier sentences. Each /h-d/ word was produced both with and without emphatic stress. Tiffany notes a shrinkage of the overall acoustic vowel space from the isolated to the stressed to the unstressed condition: “vowels seem to move toward a neutral, or at least central point on the vowel diagram as they lose energy in context” (p. 314).

Lindblom (1963) rejects the notion of a separate specification for stressed and unstressed vowels and claims that the reduced quality observed for unstressed vowels is a consequence of their shorter duration and of mechano-inertial effects. Lindblom examined the effects of changes in sentence stress and speech rate on the durations and formant frequencies of the short Swedish vowels /ɪ, e, ʏ, æ, a, ɵ, ɔ, ʊ/ in /b-b/, /d-d/ and /g-g/ contexts. For the stressed/unstressed comparison, the CVC syllables were embedded in a carrier phrase which was varied in terms of word order and rhythm. The subject was asked "to project the rhythmic patterns of the carrier sentences onto a basic periodic beat played through headphones" (p. 1774). For the slow/fast rate comparison, the subject produced stressed repetitions of the CVC syllables in time with a beat also indicated by periodic clicks over headphones, the rate of which was varied in steps from 0.5 to 6.0 cps.

Lindblom notes a similar relationship between vowel pairs produced at different rates of speech as between sententially stressed and unstressed pairs of vowels. Unstressed vowels and vowels produced at a fast rate of speech were both shorter in duration and reduced in quality compared with their stressed or more slowly produced counterparts. On the basis of these results, Lindblom argues that reduction in vowel quality is a consequence of the decrease in duration that accompanies a decrease in stress or an increase in rate and is caused by temporal overlap in the timing of motor commands to the articulators. The speaker's intention underlying the pronunciation of a given vowel is always the same (that is, the speaker always "aims" to produce a full, maximally distinct vowel) but, with an increase in rate or a decrease in stress, the articulators do not have time to reach the target configuration specified by one set of motor commands before they have to start moving in response to the next set of commands. Target undershoot at the articulatory level results in corresponding undershoot in the formant frequencies relative to the ideal spectral target. While undershoot may result in more central formant values, Lindblom argues that it may also result in values which are closer to the consonantal locus values and therefore further from the neutral or central position. In contrast to the traditional view, he thus equates reduction with contextual assimilation and not with articulatory/acoustic centralisation per se.

In subsequent studies following Lindblom (1963), it has been shown that vowel quality may vary independently of vowel duration. For example, a decrease in duration for unstressed vowels relative to the corresponding stressed vowels is not always accompanied by a reduction in vowel quality (Den Os, 1988; Fourakis, 1991). Conversely, a reduction in vowel quality may occur despite relatively long vowel durations (Nord, 1986). Nord's observations are based on lexically unstressed vowels produced in utterance-final position where relatively long durations were maintained despite reduced stress through the effects of pre-pausal lengthening.

Evidence that variation in stress and rate involve different transformations of motor activity, producing differential articulatory and acoustic effects (Verbrugge & Shankweiler, 1977; Gay, 1978; Tuller, Harris & Kelso, 1982; Engstrand, 1988) also argues against Lindblom's (1963) proposal that duration is the chief determinant of reduction. In an electromyographic study of lexically stressed and unstressed pairs of vowels produced at two different speaking rates, Tuller, Harris & Kelso (1982) found that an increase in stress was always accompanied by an increase in both the duration and peak amplitude of activity. With an increase in rate, however, three different patterns of muscle activity were observed, none of which involved decreases in both duration and peak amplitude. Tuller, Kelso & Harris' observations are based on four-syllable nonsense words of the form /əpipipə, əpipibə, əpepepə, əpepebə/ with stress placed on either of the two medial syllables, and on two-syllable nonsense words of the form /pipip, pɑpɑp, pipɑp, pɑpip/ with stress placed on either the first or second syllable. The four-syllable nonsense words were read from a list while the two-syllable nonsense words were produced in a carrier phrase.

At the acoustic level, Verbrugge & Shankweiler (1977) report larger differences in the spectral characteristics of vowels produced in /p-p/ syllables as a function of sentence stress than as a function of speech rate. Similarly, in a study of CVCVC utterances of the form /kipap'/, /ki'pap/, /kapip'/, /ka'pip/ embedded in the carrier phrase "It's a ___ again", Gay (1978) also notes that unstressed vowels, even when of a comparable duration, are "considerably reduced in F0, and somewhat reduced in overall amplitude and vowel colour with respect to their fast stressed counterparts" (p.228). Engstrand (1988) also reports an influence of stress but no influence of rate on the articulatory and acoustic properties of vowels in his data. Specifically, stressed vowels were found to be characterised by narrower oral tract constrictions and more extreme formant values than their unstressed counterparts. Engstrand examined the vowels /i, ɑ, u/ embedded in the carrier phrase "saga p-p igen" ("to say p_p again"). Stress was systematically varied between the VCV syllable and the final syllable in the carrier phrase by instructing the subjects to pronounce them in a "topic" versus "comment" mode.

Tuller, Harris & Kelso (1982) attribute the lack of a significant rate effect in their data to the fact that speakers have different strategies for effecting changes in rate. This may also account for the differential effects of rate and stress observed by Gay (1978) and Engstrand (1988). Kuehn & Moll (1976) have shown that while some speakers achieve an increase in speech rate by decreasing articulatory displacement (with a corresponding decrease in articulatory velocity), others increase rate by increasing articulatory velocity (with a corresponding decrease in the amount of undershoot). In addition to simply speeding up articulatory movements, speakers are also able to re-order the relative timing of successive gestures and so achieve an increase in speech rate by increasing the degree of temporal overlap between individual segments: ".. articulator movement toward the vowel target can begin earlier in time, that is, closer to the time of the initial consonant closure..." (Gay, 1978, p. 226). Gay (1978) and Engstrand (1988) both report increased coarticulation between segments with an increase in speech rate.

The fact that speakers appear to have some control over coarticulatory parameters indicates that vowel reduction is not simply an automatic consequence of a decrease in duration and of mechano-inertial effects. In recognition of this, Lindblom (1983, 1988, 1990, 1992) revises his original undershoot hypothesis to include a compensatory mechanism whereby speakers are able to override duration-dependent undershoot through increased vocal effort and more forceful opening and closing gestures. However, he continues to hold to his original conviction that reduction reflects coarticulatory effects rather than the intention to produce a centralised vowel: "Reduction processes can be seen as contextual assimilations durationally induced, but within certain limits, speakers appear capable of controlling the precise degree of reduction" (Lindblom et al., 1992, p. 362).

An alternative account of vowel reduction which reconciles elements of both the traditional view and of Lindblom's analysis, is an account in terms of phonetic underspecification. In line with Keating's (1988, 1990) phonetic underspecification hypothesis, it is suggested here that unstressed vowels are inherently less fully or narrowly specified than their stressed counterparts. Thus, in accordance with the traditional concept, generally less extreme articulatory and less distinctive spectral configurations are predicted for unstressed compared with stressed vowels. However, in contrast to the traditional account and in keeping with Lindblom's analysis, vowel reduction is not viewed as an independent process of articulatory and acoustic centralisation towards the centre of the vowel space. Rather, in view of the evidence presented in Chapter 5 that schwa is unspecified for F1 and F2 (see also Van Bergem, 1994 and Kondo, 1995) and can therefore occupy almost any position in the vowel space depending on context, it is predicted that reduced vocal effort and less divergence from neutral in the case of unstressed vowels will result in greater contextual assimilation. (Here 'neutral' is used in an abstract sense insofar as it represents a notional starting point. While this may correspond to the neutral rest configuration in some instances, in many cases, it will represent an intermediate point on the trajectory between adjacent context segments.)

The principal way in which this analysis differs from Lindblom's revised undershoot hypothesis is with respect to the nature of the proposed underlying representations.

Lindblom posits an invariant target for both stressed and unstressed vowels with varying degrees of target undershoot depending upon the degree of vocal effort employed to override coarticulatory effects and with presumably less effort made in the case of unstressed vowels. The phonetic underspecification account assumes a separate specification for stressed and unstressed vowels which implies reduced vocal effort. In terms of the phonetic consequences, both models make the same prediction, i.e. increased contextual assimilation for unstressed vowels.

Since the majority of studies following Lindblom (1963) which investigate the effects of stress on vowel quality focus on the proposed relationship between reduction and duration, relatively little attention has been paid to the question of how stress and context effects interact to influence vowel quality. In general, the results of these studies confirm longer durations, more extensive spatial movements and more extreme formant frequency values for stressed compared with unstressed vowels resulting in a global increase in the size of the vowel space for stressed vowels relative to unstressed vowels. Comparatively little evidence of differences in the extent of local context effects across stress conditions has been reported. However, this may be attributed to the fact that, with few exceptions, vowels are examined in CVC or VCV nonsense words using a restricted set of vowels and consonants. Given that vowel reduction forms part of the speaker's stylistic intent, the artificial nature of the speech material employed in these studies is also likely to inhibit effects.

Notwithstanding the need for more extensive research, there is some evidence in the literature that unstressed vowels are more sensitive to context effects than stressed vowels. For example, Nord (1974) reports greater coarticulation for the Swedish vowels /a/ and /e/ in unstressed compared with in stressed position. The vowels were examined in a /s-l/ frame either followed or preceded by an additional VC or CV syllable. Stress contrasts were obtained by producing the words with stress on either the initial or final syllable. In addition to consonant-to-vowel effects, unstressed /a/ was also found to be more sensitive to the influence of a following stressed vowel (see also Nord, 1986).

Fowler (1981) also reports larger vowel-to-vowel coarticulatory effects for unstressed /ʌ/ than for stressed /ʌ/ in $V^1bʌbV^3$ sequences. The tri-syllabic sequences were read from two lists. In the first list, V^1 and V^3 were stressed and the medial vowel unstressed (as in "misinform"). In the second list the reverse pattern obtained whereby V^1 and V^3 were unstressed while the medial vowel was stressed (as in "deliver"). Fowler explains her results in terms of the co-production of stressed and unstressed vowels:

> coarticulatory effects of stressed on unstressed vowels may signify that the talker subsumes the production of unstressed vowels within the domain of the production of a preceding vowel, and, to a lesser extent, relative to the beginning phases of a following vowel (p. 137).

Interestingly, she observed contextual variation only in the case of F2. However, while there was no influence of adjacent vowels on the F1 values for /ʌ/, there was a significant influence of stress, with the stressed tokens showing higher F1 values than the unstressed tokens. Similar global effects on F1 as a function of stress have also been noted by Lisker (1984), Summers (1987) and Van Son (1993).

In an acoustic study of rate and stress effects, Fourakis (1991) examines the nine non-retroflex American English monopthongs in /b-d/ and /h-d/ contexts embedded in two carrier phrases: "I will say kay___ again" and "I will say KAY ___ again". In the first of these, subjects were instructed to pronounce the disyllabic compound with main stress on the second syllable while in the second, subjects were instructed to pronounce the compound with main stress on 'KAY'. Fourakis found that context exerted a greater influence on the spectral characteristics of vowels at the individual token level than either rate or stress, despite the fact that only two consonsonantal frames were used and that these, furthermore, are assumed to interfere minimally with vowel articulation. In terms of the overall vowel space, the reverse trend was noted. While no variation in the size of the vowel space occurred as a function of context, a reduction of 30% was observed in the fast, unstressed condition relative to the slow, stressed condition. Because the shrinkage of the vowel space did not result in a loss of acoustic contrast between vowels, Fourakis concludes that phonetic vowel reduction is not conditioned by rate or stress. However, what Fourakis does not take into consideration and which is implicit in the undershoot hypothesis, is that the effects of stress and context may be additive. The overall shrinkage in the vowel space he observes for fast, unstressed vowels relative to slow, stressed vowels is consistent with the view that unstressed vowels are characterised by less extreme articulatory and spectral configurations. The reason this was not accompanied by a loss of acoustic contrast between individual vowels may be attributable to the fact that only two contexts were examined. While formant values may display a shift in the /b-d/ frame relative to the /h-d/ frame, the direction of the shift is likely to be the same for all vowels in which case the relative distance between vowels is unlikely to be affected.

In addition to the questions of whether and how far stress and context effects interact to influence vowel quality, there is also the question of whether stress affects all vowels to the same degree or whether some vowels show relatively greater shifts in value as a function of stress than others. Again, there is little comparative data in the literature since the majority of studies which investigate stress effects on vowel quality examine only a restricted number of vowels.

Assuming that degree of reduction is inversely related to duration, relatively greater reduction in quality and consequently a greater overall difference in amount of coarticulation between stressed and unstressed conditions might be expected for the long vowels /ɑ/, /ɜ/, /ɔ/ and /a/ compared with the short vowels /ɪ, ɛ, ʌ, ʊ/. This prediction is made in accordance with Klatt's (1973, 1975, 1976) proposal that there is an "incompressibility" limit on segmental durations such that inherently short segments are more resistant to shortening as a function of factors such as stress and rate than inherently long segments.

However, while the long vowels may show greater absolute shortening than the short vowels, the evidence also indicates that the percentage changes in duration remain the same across vowels (Peterson & Lehiste, 1960; Klatt, 1973). Thus, given that the relative degree of shortening is the same for all vowels, the relative degree of coarticulation shown by different vowels might also be expected to remain constant across stress conditions. In other words, the pattern across vowels with respect to which vowels show most sensitivity to context is not expected to change as a function of stress, only the relative magnitude of coarticulatory effects within each vowel category.

In the following analysis, the role of sentence stress and, more importantly, the interaction of sentence stress and context in determining vowel quality is examined. Sententially stressed and unstressed vowel tokens are compared in terms of the overall variability and the patterns of variability they exhibit with respect to the following questions:

(a) How variable are unstressed vowels in comparison with stressed vowels?
(b) Are unstressed vowels more sensitive to context than stressed vowels?
(c) Do some vowels show relatively greater shifts in vowel quality as a function of sentence stress than others?
(c) Is the hierarchy of vowel robustness evident in Chapter 5 preserved across stress conditions?

The role of sentence stress in determining vowel quality is investigated in the light of the hypothesis that unstressed vowels are more susceptible to coarticulatory effects than stressed vowels because they are less narrowly specified. If this analysis is correct, unstressed vowels are expected to show greater overall variability and greater overall context-dependency than their stressed counterparts.

6.1 Data

All vowel tokens were hand-labelled for sentence stress using both aural and visual cues. Perceptual judgements were supplemented by reference to visual displays of vowel duration, vowel amplitude and pitch movement. Three levels of sentence stress were assigned: 'nuclear accented', 'non-nuclear accented' and 'unaccented'. The first level of stress was assigned to lexically stressed vowels which were judged to bear nuclear accents, the second level to all other lexically stressed vowels also judged to bear sentence stress. Unaccented vowels represent all vowels that were unaccented but which were nevertheless considered to be unreduced.

Vowels were marked for sentence stress as opposed to lexical stress on the grounds that while all lexically stressed vowels carry the potential for sentence stress, they do not always receive it when produced in fluent speech. As Lehiste (1970) writes:

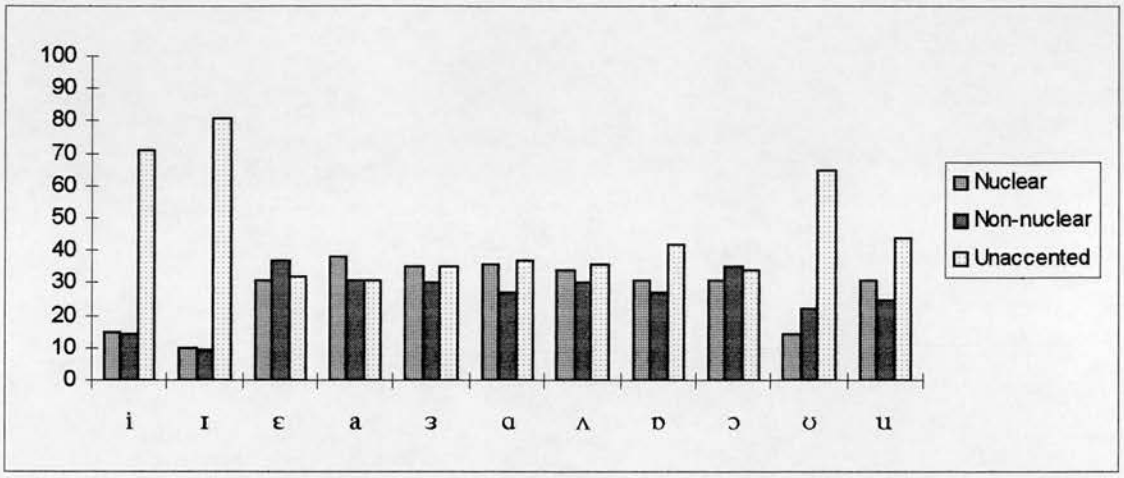
> It is probable that word-level stress is, in a very real sense, an abstract quality: a potential for being stressed. Word-level stress is the capacity of a syllable within a word to receive sentence stress when a word is realised as part of the sentence. The degrees of stress of other syllables within the word are usually predictable by rules and are therefore non-contrastive. The fact that not all syllables that are perceived as stressed are associated with peaks of subglottal pressure supports the idea that what is realised phonetically is sentence-level stress rather than word-level stress. In other words, our knowledge of the structure of the language informs us which syllables have the potential for being stressed; we 'hear' the underlying phonological forms. (p. 237)

The number of tokens within each stress condition (henceforth referred to as 'accent condition') for each vowel is given in Table 6.1. The percentage of nuclear and non-nuclear accented compared with unaccented tokens is also shown graphically in Figure 6.1. The comparatively high ratio of unaccented to accented tokens evident for /ɪ/, /ʊ/, and /i/, reflects the relatively high occurrence of these vowels in grammatical compared with in lexical items.

The preceding and following contexts for each vowel in each accent condition are listed in Tables 6.17:18. The proportional distribution of contexts is also shown in Figures 6.16:26. Place and manner categories are abbreviated as before (see section 5.1).

Henceforth all references to 'stress' are references to sentential stress as opposed to lexical stress.

Figure 6.1: Percentage of nuclear and non-nuclear accented and unaccented tokens within each vowel category.



6.2 Analysis

The mean values and standard deviations for duration and for F1 and F2 midpoint values for each vowel for each accent condition are shown in Tables 6.1:3. To test for global effects of sentence stress on each parameter, one-way analyses of variance and Tukey multiple pairwise comparison of means tests were performed (see section 5.3). The results of the analyses of variance are presented in Tables 6.4:6.

6.2.1 Duration

The mean durations for each vowel, for each accent condition are also shown graphically in Figure 6.2. In addition, the degree of shortening for each vowel across accent conditions is plotted in Figure 6.3. This was calculated by dividing the difference in mean duration between the two accent conditions being compared (i.e. nuclear accented versus unaccented) by whichever mean duration is the highest and multiplying the result by 100.

In accordance with the literature, the present data shows a significant main effect of sentence stress on vowel duration for all vowels, with longer durations accompanying increased stress. With the exception of /ʊ/, nuclear accented tokens are significantly longer in duration for all vowels than either the corresponding non-nuclear accented or unaccented tokens. The greatest difference in mean duration between nuclear accented and unaccented tokens occurs for /u/ and /i/ (40% and 34% respectively) followed by /ʊ/ and /ɪ/ (27%). The remaining vowels show a comparable degree of shortening. The greatest difference in mean duration between nuclear and non-nuclear accented tokens occurs for /u/, /ɜ/, /ɔ/ and /ɪ/. In the case of /ʊ/, there is no difference in duration between nuclear and non-nuclear accented tokens. For the vowels /i, ɪ, ɒ, ʊ, u/, there is also a significant difference in the duration of non-nuclear accented compared with unaccented tokens, the accented tokens predictably showing longer durations. Most difference occurs for /ʊ/ (25%) and /ɒ/ (23%). The remaining vowels /ɛ, a, ɑ, ʌ, ɔ, ɜ/ show no significant durational difference between non-nuclear accented and unaccented tokens.

Individual differences in the relative amount of shortening with a decrease in stress predictably result in differences in the durational hierarchy across accent conditions. In the accented (combined nuclear and non-nuclear) condition, /i/ and /u/ tokens show comparable durations to /a/ tokens and significantly longer durations than the short vowels /ɛ/ and /ʌ/. In the case of /i/, the difference is significant at $p < .001$. The difference in duration between /u/ and /ɛ/ is significant at $p < .01$ and the difference in duration between /u/ and /ʌ/, significant at $p < .05$. Accented /u/ tokens also show significantly longer durations than accented /ɒ/ tokens ($p < .05$). In the unaccented condition, however, durations for /i/ and /u/ are significantly shorter than durations for /a/ ($p < .001$) and comparable to those for /ɒ/, /ɛ/ and /ʌ/.

Other differences include significantly longer durations for /ɑ/ in the accented condition compared with /ɜ/ ($p < .01$) and /ɔ/ ($p < .001$). Unaccented /ɑ/, /ɜ/ and /ɔ/ tokens show no significant durational difference. Accented /ɒ/ tokens also show significantly longer durations than accented /ɛ/ ($p < .05$) and /ʌ/ tokens ($p < .01$). However, there is no durational difference between unaccented /ɒ/, /ɛ/ and /ʌ/ tokens.

Nuclear accented /u/ and /i/ tokens show significantly longer durations than nuclear accented /ɛ/ and /ʌ/ tokens ($p < .001$). However, in the non-nuclear accented condition, there is no significant difference in duration for /i/ tokens compared with /ɒ/ tokens or for /u/ tokens compared with /ɒ, ɛ, ʌ/ tokens.

Overall, there is, predictably, greater durational overlap for unaccented tokens than for accented vowel tokens. Vowels also show greater overlap in the non-nuclear accented compared with in the nuclear accented condition. However, differences in the ranking of vowels across accent conditions chiefly concerns the vowels of intermediate duration /i, u, ɒ, ɛ, ʌ/. The durational hierarchy is preserved with respect to which vowels show the longest and which show the shortest durations. The tense vowels /ɑ/, /ɜ/ and /ɔ/ show the longest durations in all three conditions while /ʊ/ and /ɪ/ consistently show the shortest durations. The long, lax vowel /a/ is also characterised by longer durations within each accent condition than the short vowels /ɒ/, /ɛ/ and /ʌ/.

6.1.2 F2

There is a significant main effect of sentence stress on F2 midpoint values in the case of the vowels /i, ɪ, ʌ, ɒ, ɔ, ʊ, u/. All pairwise comparisons between nuclear accented, non-nuclear accented and unaccented values are significant for /i/ at $p < .01$, with nuclear accented tokens displaying the highest midpoint values and unaccented tokens, the lowest midpoint values. In the case of /ɔ/, the difference in F2 value between nuclear accented and unaccented tokens is significant at $p < .001$, lower values being obtained for the accented tokens. Midpoint values also tend to be lower for nuclear compared with non-nuclear accented /ɔ/ tokens although the difference only reaches significance at $p < .1$. For each of the other vowels, there is no distinction between nuclear and non-nuclear accented tokens. In the case of /ʌ/, F2 values are generally lower for accented than for unaccented tokens. However, the difference in value only attains significance for non-nuclear accented compared with unaccented tokens ($p < .01$). For /ɒ/ and /u/ there is a small difference in F2 midpoint value between nuclear accented and unaccented tokens ($p < .1$). The difference between non-nuclear accented and unaccented tokens is also significant at $p < .1$ in the case of /ɒ/ and at $p < .01$ in the case of /u/. Unaccented /ɒ/ tokens display higher values than the corresponding accented tokens while unaccented /u/ tokens display lower values than their accented counterparts. There is also a tendency for values to be lower for unaccented compared with nuclear accented /ɪ/ tokens ($p < .1$). There is no significant difference in F2 value between non-nuclear accented and unaccented tokens in the case of /ɪ, ɛ, a, ɜ, ɑ, ʌ, ɔ/.

6.2.3 F1

Sentence stress has a greater overall effect on F1 than on F2. The only vowel which does not show a significant difference in F1 midpoint value as a function of stress is /ɔ/. In the case of /ʌ/, all pairwise differences are significant at either $p < .001$ (nuclear accented versus unaccented) or at $p < .05$. There is no significant difference between non-nuclear accented and unaccented tokens for /i, ɪ, ɛ, ɑ, ɒ/. However, in each case, there is a significant difference in value between nuclear accented and unaccented tokens. Nuclear accented /i/ tokens are characterised by lower F1 values than the corresponding unaccented tokens. For each of the other vowels, nuclear accented tokens display higher values than their unaccented counterparts. These vowels also display a significant difference in value between nuclear and non-nuclear accented tokens, although in the case of /i/, the difference is only significant at $p < .1$. For /a/ and /ʊ/, nuclear and non-nuclear accented tokens are characterised by higher values than unaccented tokens while for /u/ unaccented tokens display higher values than the corresponding nuclear and non-nuclear accented tokens. There is, however, no difference in value between tokens in the two accented conditions. In the case of /ɜ/, non-nuclear accented tokens display lower values than nuclear accented tokens ($p < .001$). There is no difference in value between either nuclear accented or non-nuclear accented /ɜ/ tokens compared with unaccented /ɜ/ tokens.

6.3 Distribution of formant values

In order to compare accented and unaccented vowel tokens with respect to the overall variability they exhibit, formant values, sampled at the durational vowel midpoint and converted to a Mel scale, were plotted in F1-F2 acoustic space. Ellipses were drawn to two standard deviations round the data points. The area of the ellipse thus provides an index of the combined variance along F1 and F2 for each vowel. Figures 6.4:6.10 show the distribution of values for tense and lax vowels plotted separately as a function of accent condition. In view of its long inherent duration and acoustic stability, the long, lax vowel /a/ is included in the tense vowel set. In Figures 6.4 and 6.5, nuclear and non-nuclear accented tokens are combined. In each figure, vowel tokens are pooled across all contexts. Comparative measures of the relative size of each ellipse are given in Table 6.5. These represent the overall area of each ellipse scaled down by a factor of four. The degree of shrinkage in ellipse size as a function of bearing a nuclear compared with a non-nuclear accent and of bearing an accent (nuclear and non-nuclear combined) compared with no accent is also shown. This was obtained by expressing the smaller ellipse for each pairwise comparison as a percentage of the area of the larger ellipse and subtracting it from 100. The reduction in overall variability for each vowel as a function of stress is also plotted in Figure 6.13. The coefficient of variance values for each formant dimension are given in Table 6.6. They are also represented graphically in Figure 6.13.

6.3.1 Size and orientation of ellipses

6.3.1.1 Size of overall vowel space

In both tense and lax vowel sets, accented tokens display a tighter clustering of values than the corresponding unaccented tokens. With the exception of /ɜ/, the tense vowels are also characterised by more peripheral values when accented than when unaccented. As the results of the pairwise comparison of means tests demonstrate (see section 6.1.3:4), accented /i/ and /u/ tokens generally display higher F2 and lower F1 values than unaccented /i/ and /u/ tokens. Accented /ɔ/ tokens are generally characterised by lower F2 values than their unaccented counterparts while /a/ and /ɑ/ show generally higher F1 values when accented than when unaccented. In the case of /i/, /u/ and /ɑ/, the difference in F1 reflects a difference between the nuclear accented and unaccented tokens. There is no difference between non-nuclear accented and unaccented tokens for these vowels. In the case of /a/, F1 is higher for both nuclear and non-nuclear accented compared with unaccented tokens.

The accented /ɜ/ tokens occupy an area at the centre of the distribution of the unaccented /ɜ/ tokens. There is no difference in formant frequency value between the combined accented and unaccented tokens. However, nuclear accented tokens are characterised by higher F1 values than unaccented tokens.

In contrast to the tense vowels, the overall vowel space for lax vowels is smaller for the accented compared with the unaccented tokens, the accented tokens being located more centrally within the distribution of the corresponding unaccented tokens. However, there is a tendency for /ʌ/ and /ɒ/ to show lower F2 values and higher F1 values when accented than when unaccented. /ɪ/ and /ɛ/ also show higher F1 values in stressed compared with in unstressed position.

6.3.1.2 Direction of spread

Ellipses show a similar orientation across accent conditions in the case of the vowels /i, ɑ, ɪ, ɛ, u, ʊ, ɔ/, with /i/, /ɑ/ and /ɪ/ displaying greater variability along F1 than along F2 and the remaining vowels displaying comparatively greater F2 variability. Unaccented /a/ tokens are more variable along F1 than along F2 while nuclear and non-nuclear accented /a/ tokens show greater F2 variability. In the case of /ʌ/, unaccented tokens show greater variability along F2 while nuclear and non-nuclear accented tokens are more variable along F1. For /ɜ/ and /ɒ/, the nuclear accented and unaccented tokens pattern together showing greater variability along F1 in the case of /ɜ/ and greater variability along F2 in the case of /ɒ/.

6.3.1.3 Overall variability

With the exception of /ɜ/, unaccented tokens show greater variability than either nuclear or non-nuclear accented tokens for all vowels. In the case of /ɜ/, non-nuclear accented tokens are slightly more variable than unaccented tokens. Non-nuclear accented tokens also show a larger combined variance than nuclear accented tokens in all cases except /i, ɛ, a, u/. For /i/ and /ɛ/, the difference in ellipse size between the two accented conditions is relatively small. In the case of /a/, the difference is largely accounted for by greater F1 variability, and for /u/, greater F2 variability, for nuclear compared with non-nuclear accented tokens.

As Figure 6.13 demonstrates, /ʌ/ and /ʊ/ show the greatest difference in overall variability between the combined accented and the unaccented condition, followed by the tense vowels /ɜ/, /ɔ/ and /ɑ/. The large difference in overall variability for combined accented compared with unaccented /ɜ/, /ɔ/ and /ɑ/ tokens is largely accounted for by the difference in variability between nuclear accented and unaccented tokens. In the comparison between non-nuclear accented and unaccented tokens, /ɒ/ and /u/ show the greatest reduction in overall variability followed by /ʌ/ and /ʊ/. /ɜ/ shows the least difference in overall variability between the non-nuclear accented and unaccented condition after /ɛ/. /ɛ/ shows the least difference in ellipse size across accent conditions.

Despite individual differences between vowels with respect to the amount of reduction in variability they show with an increase in stress, with the exception of /ɯ/, the tense vowels as a set display lower overall variability within each accent condition than the lax vowels. /ɜ/ displays the lowest combined variance in the nuclear accented and unaccented conditions, followed by /ɑ/. In the non-nuclear accent condition, /ɜ/ displays the lowest combined variance after /a/ and /i/. In the unaccented condition, /ʊ/ and /ɒ/ display the highest overall variability following /ɪ/. /ɪ/ shows the greatest overall variability within each accent condition.

6.3.2 Degree of overlap

Percentage figures for the degree of overlap between vowel categories within the combined accented and the unaccented conditions are given in Tables 6.9:10. For each pairwise comparison, these represent the number of tokens for each pair member that fall within the ellipse of the other pair member, expressed as a percentage of the total number of tokens for both pair members. The degree of overlap between each full vowel and schwa for each accent condition is also given.

As Tables 6.7:8 demonstrate, the tighter clustering of values observed for accented compared with unaccented tokens predictably results in greater overall separability between individual vowel categories. For both tense and lax vowels, the accented tokens are more distinct as a set than the unaccented tokens. In the case of the tense vowels, the only overlap among accented tokens occurs between /i/ and /ɯ/ (5%) and between /a/ and /ɑ/ (1%). The degree of overlap between these vowels increases in the unaccented condition to 33% and 2% respectively. The number of unaccented vowel pairs that overlap also increases to include /ɯ/-/ɜ/, /ɯ/-/ɔ/, /ɜ/-/ɑ/, /ɜ/-/ɔ/ and /ɔ/-/ɑ/.

The difference in the degree of overlap across accent conditions is most marked in the case of the lax vowels. For example, the high percentage of overlap between /ʌ/ and /ɒ/ tokens in the unaccented condition (60%) reduces to 24% in the accented condition. Large differences in amount of overlap as a function of stress are also evident for the following pairs: /ɪ/-/ʊ/, /ɛ/-/ʌ/, /ɛ/-/ɒ/, /ɒ/-/ʊ/. The degree of overlap between lax vowels and schwa is also predictably less for accented tokens compared with unaccented tokens.

The tighter clustering of accented vowel tokens within both the tense and lax sets also results in a lesser degree of overlap between tense-lax pairs of vowels. The greatest difference in amount of overlap as a function of stress occurs for /ʊ/ and /ʊ/, /ɛ/ and /ɜ/ and /ʌ/ and /ɜ/. Stress makes little difference in the case of the tense-lax pair /ɑ/-/ʌ/. This reflects the fact that the generally high degree of overlap between /ɑ/ and /ʌ/ is attributable to the inherent spectral similarity of these vowels rather than to contextual variability.

6.4 Overall context-dependency

The overall degree of context-dependency shown by accented and unaccented tokens for each vowel category was assessed in a series of multiple regression analyses. As in the case of the regression analyses in Chapter 5, the dependent variable was F1 or F2 sampled at the durational vowel midpoint. The predictor variables were preceding context (lhs), following context (rhs) and vowel duration (dur). Results are presented in Tables 6.9:13. The proportion of explained variance for each vowel in each accent condition is also represented graphically in Figure 6.15:16.

6.4.1 F2

6.4.1.1 Total proportion of explained variance

In the combined accented versus unaccented comparison, a significant difference in the total proportion of explained F2 variance for accented compared with unaccented tokens is obtained in the case of all vowels except /ɜ/, /ɛ/ and /ɑ/. The amount of difference varies across vowels and is greatest, in descending order, for: /ɒ/ and /u/ (25%), /ɔ/ (23%), /a/ (19%), /ʌ/ (11%). The front vowels /i/ and /ɪ/ show the least difference across accent conditions (5%). Consequently, the rank ordering of vowels with respect to the total amount of context-dependency shown, differs for the accented compared with the unaccented condition. In the case of /u/, /ɒ/ and /ʌ/, the difference is in the expected direction, resulting in unaccented tokens occupying a relatively higher position in the ranking of vowels from most to least context-dependent than the corresponding accented tokens. In the case of /ɔ/, however, the opposite trend is apparent with accented tokens displaying a higher R^2 value than their unaccented counterparts. (/ɜ/ also shows a higher R^2 value for accented compared with unaccented tokens but the difference is not significant.) The relatively large difference in amount of context-dependency shown by /a/ has little effect on its rank position. For both accent conditions, it numbers among the three vowels which show the least overall context-dependency.

For the remaining vowels, there is comparatively little difference in the relative amount of context-dependency shown by each across accent conditions. In all cases, unaccented tokens show greater context-dependency than accented tokens. The lax vowels /ɪ/, /ʊ/ and /ɛ/ and the tense vowel /u/ display the highest R^2 values and /ɑ/, the lowest R^2 value in both accent conditions.

The only vowels to show a significant difference in amount of context-dependency for nuclear accented compared with non-nuclear accented tokens are /ɪ/, /ɔ/ and /ɒ/. In each case, the difference is in the expected direction with non-nuclear accented tokens showing greater context-dependency than nuclear accented tokens. For /ɔ/ and /ɒ/, the difference serves to alter their position in the rank ordering of vowels with respect to which shows the greatest context dependency. In the case of /ɪ/, however, there is no change in its ranking. In both nuclear and non-nuclear accented conditions, it shows the highest levels of context-dependency after /ʊ/.

The results for non-nuclear accented compared with unaccented tokens shows that the difference between the combined accented /a/ tokens and their unaccented counterparts is due to a difference between nuclear accented and unaccented tokens. There is no significant difference for this vowel between the non-nuclear accented and the unaccented condition. Similarly, there is no significant effect of bearing a non-nuclear accent compared with no accent in the case of /ɑ/. Otherwise, all vowels display significantly higher levels of context-dependency for unaccented compared with non-nuclear accented tokens. All vowels with the exception of /ɛ/, /ɜ/ and /ɔ/, show a significant difference in context-dependency between nuclear accented and unaccented tokens.

6.4.1.2 Directionality of effects

In the combined nuclear and non-nuclear accented group, following context makes no significant contribution to the F2 prediction for /ɑ/, /ɒ/ or /ɔ/. In the case of /ɪ/ and /ɛ/, it accounts for a higher proportion of the variance than preceding context. There is no difference in the directionality of effects for /i/. In all other cases, carryover effects exceed anticipatory effects although for /a/ the difference in directionality is marginal.

In the unaccented group, /u/ and /ʊ/ display relatively greater anticipatory than carryover coarticulation. The directionality of effects is comparable for /ɛ/, otherwise carryover effects exceed anticipatory effects. Again, there is no significant contribution of following context for /ɑ/ or /ɔ/. With the exception of /ɛ/, /ɒ/ and /ʊ/, the difference in the relative strength of anticipatory compared with carryover effects is most marked in the unaccented condition.

The directionality of effects for nuclear accented compared with non-nuclear accented tokens show a similar pattern except that following context accounts for a relatively greater amount of the variance in /i/, /ɛ/ and /u/ in the nuclear accented than in the non-nuclear accented condition.

6.4.1.3 Duration

Among the combined nuclear and non-nuclear accented tokens, duration uniquely accounts for the highest proportion of F2 variance in the case of /i/ and accounts for more of the F2 variance than following context in the case of /ɔ/ and /ɒ/. It also makes a significant contribution to the prediction for /ɪ/, /ɛ/, /a/ and /ʊ/. It has no significant effect on the F2 equation for /ɜ/, /ɑ/, /ʌ/ or /u/.

The relative importance of duration to the prediction for /i/ diminishes in the unaccented condition and is displaced by preceding context. However, it continues to account for a significant proportion of the F2 variance and a greater proportion of the variance than is uniquely accounted for by following context. Duration continues to make a greater contribution to the variance in F2 than following context for both /ɔ/ and /ɒ/. It also uniquely accounts for more variance than following context in the case of /ɑ/. It continues to make a significant contribution to the prediction for /ɪ/, /ɛ/ and /ʊ/ although this is smaller than either the effects of preceding or following context. There is also a significant effect of duration for /ʌ/ which was not evident in the accented condition. It makes no significant contribution to the prediction in the case of /ɜ/, /a/ or /u/.

The relative weighting of the contribution made by duration varies across the two accented conditions. The significant effects of duration observed for /ɛ/, /a/ and /ʊ/ tokens in the combined accented condition reflect significant effects for non-nuclear accented tokens only. There is no significant effect of duration for nuclear accented /ɛ/, /a/ or /ʊ/ tokens. The relative influence of duration on the prediction for /i/ is also diminished in the nuclear accented relative to the non-nuclear accented condition although it continues to make a significant contribution. In addition, there is also a significant effect of duration for /ʌ/ in the non-nuclear accented condition which is not evident for the corresponding nuclear accented tokens.

6.4.1.4 Interaction of context effects

For both the accented and unaccented tokens, a significant interaction of context effects is only evident in the case of /i/ and /ɪ/. Accented /i/ tokens, show a small correlation between onset and offset value, r=.031 which is significant at $p < .001$. The corresponding unaccented tokens show a stronger correlation, r=.168 which is significant at $p < .0001$. The difference in the strength of the correlation for accented and unaccented tokens is significant at $F(2, 660) = 3.36$, $p < .035$. In the case of /ɪ/, accented tokens show a correlation of r=.149 ($p < .02$) which increases in the unaccented condition to r=.441 ($p < .0001$). The difference in the strength of the correlation across accent conditions is also significant at $F(2, 1354) = 16.41$; $p < .0001$.

6.4.2 F1

6.4.2.1 Total proportion of explained variance

There is no significant difference in the amount of explained F1 variance for combined accented compared with unaccented tokens in the case of /ɛ, ɜ, ɑ, ɔ/. All other vowels, with the exception of /ʊ/, show a significantly higher proportion of explained F1 variance for unaccented compared with accented vowel tokens. The difference in the overall R^2 value across accent conditions is greatest in the case of /u/ (50%), /a/ (34%), /ɒ/ (29%) and /ɪ/ (24%). The difference is less although still significant for /i/ (17%) and /ʌ/ (9%).

In the case of /ʊ/, accented tokens display a higher R^2 value than their unaccented counterparts.

For /u/, /ɪ/ and /ʌ/, the difference between accented and unaccented tokens reflects a difference between the nuclear accented and unaccented conditions. For these vowels, there is no difference in overall context-dependency between non-nuclear accented and unaccented tokens. Similarly, the absence of a significant effect for accented compared with unaccented /ɑ/ tokens reflects the lack of a significant difference between non-nuclear accented and unaccented tokens since nuclear accented /ɑ/ tokens show a significantly lower level of context-dependency compared with unaccented /ɑ/ tokens.

6.4.2.2 Directionality of effects

In the combined accented group, carryover effects exceed anticipatory effects in the case of /a/, /ʌ/, /ɒ/ and /ʊ/. For the other vowels, anticipatory effects are dominant. Among the unaccented tokens, following context accounts for a higher proportion of the variance than preceding context in the case of /ʊ/ and /u/. Otherwise, carryover effects predominate.

There are some differences in the directionality of effects between the two accented conditions. Following context makes a slightly greater contribution to the variance in nuclear accented /i/ tokens than preceding context. The opposite applies in the case of non-nuclear accented /i/ tokens. There is also less difference in the directionality of effects for non-nuclear accented compared with nuclear accented /i/ tokens. Following context makes no significant contribution to the prediction for /a/ in the nuclear accented condition, however, it does have a significant effect in the non-nuclear accented condition. In the case of /ɑ/, there is no significant effect of preceding or following context for non-nuclear accented tokens. However, there is a significant carryover effect for nuclear accented tokens. Preceding and following context account for comparable amounts of variance in /ɒ/ in the nuclear accented condition although following context accounts for no variance in non-nuclear accented /ɒ/ tokens. In the case of /ʊ/, there is no significant effect of either preceding or following context for nuclear accented tokens. For non-nuclear accented tokens, however, there is a significant effect of both preceding and following context with preceding context making the greatest contribution to the variance. There are no carryover effects for /u/ within either of the accented conditions. The directionality of effects for /ʌ/ is comparable for both nuclear and non-nuclear accented tokens with preceding context accounting for slightly more of the variance than following context in the case of the non-nuclear accented tokens. These differences in the relative magnitude of carryover and anticipatory effects between accent conditions may be attributable, in part, to differences in the frequency distribution of contexts for each condition.

6.4.2.3 Duration

In the accented condition, duration makes a significant contribution to the F1 prediction for all vowels except /ɪ/, /ɑ/, /ɔ/ and /u/. It uniquely accounts for the highest proportion of F1 variance in the case of /ɜ/ and /ʌ/ and makes a greater contribution to the overall variance than following context in the case of /a/ and /ɒ/. It also uniquely accounts for more variance than preceding context in the case of /ɛ/. For the unaccented tokens, duration accounts for the highest proportion of F1 variance in /a/ and /ɜ/. It also makes a greater contribution to the F1 prediction than following context for /ɑ/ and /ʌ/. It makes no significant contribution in the case of /ɔ/, /ʊ/ or /u/.

As in the case of F2, the relative importance of duration to the equation for F1 varies across the two accented conditions. Most notably, whereas duration contributes most of all the predictors to the variance in nuclear accented /ɒ/ tokens it makes no significant contribution to the variance in non-nuclear accented /ɒ/ tokens. The relative importance of duration is also diminished in the non-nuclear accent condition relative to the nuclear accent condition in the case of /ɜ/, /ʌ/, /ɛ/ and /ʊ/. In the case of /a/, duration makes the highest contribution to the prediction for non-nuclear accented tokens. In the nuclear accented condition, it accounts for more of the variance in /a/ than following context but not more than preceding context.

6.4.2.4 Interaction of context effects

/ɪ/ is the only vowel to show a significant interaction of effects for either accented or unaccented vowel tokens. The correlation between F1 onset and offset value for unaccented tokens of $r=.403$ ($p < .0001$) is stronger than the correlation for accented tokens ($r=.161$, $p < .01$). The difference is also significant at $F(2, 1534) = 20.99$, $p < .0001$. While unaccented /i/ tokens also show a significant correlation between onset and offset value ($r=.130$, $p < .0001$), the correlation for corresponding accented tokens is not significant ($r=.107$, $p < .13$) and the difference between r values for accented and unaccented tokens also fails to attain significance, $F(2, 660) = 1.59$, $p < .203$.

6.5 Local context effects

In Chapter 5, analyses of variance and pairwise multiple comparison of means tests were performed in order to assess the separability in mean midpoint value for schwa as a function of C^1 and C^2 contexts in comparison with the full vowels and to complement the results of the regression analyses with respect to the evaluation of the directionality of effects. The high degree of separability observed for schwa and /ɪ/ as a function of both preceding and following consonantal place and manner of articulation confirmed the transitional nature of these vowels. While it is hypothesised that unaccented vowels are less narrowly specified than accented vowels, the unaccented vowel tokens in the present data are not considered to be reduced and hence are not expected to display the same trends as schwa or /ɪ/ either in terms of the magnitude or the pattern of coarticulatory effects.

Given that the higher R^2 values observed for unaccented relative to accented vowel tokens reflects greater context-dependency, unaccented tokens should show relatively greater shifts in formant frequency value as a function of individual contexts than the corresponding accented tokens. For example, lower absolute values for F2 are predicted for unaccented /i/ tokens in the context of a preceding /w/ or following /ɫ/ than for accented /i/ tokens in equivalent contexts and, similarly, higher values are predicted for unaccented /ɔ/ tokens compared with accented /ɔ/ tokens in palatal or alveolar contexts. However, whether such effects result in greater overall separability as measured in terms of the number of pairwise differences that attain significance, depends on their relative strength compared with the strength and direction of influence of other contexts. Given that, for example, in the case of /ɔ/, most contexts will exert a raising influence on the vowel midpoint values, a general upward shift in F2 for unaccented /ɔ/ tokens may actually serve to lessen the overall degree of separability in midpoint value.

Since greater sensitivity to context for unaccented vowels is expected to be manifest primarily in a greater magnitude of effects rather than in differences in the directionality of effects, the pairwise comparisons are used here to complement the regression analysis by indicating the number and type of contexts that contribute to the whole effect.

Owing to the limitations in sample size across place (and manner) categories imposed through the introduction of sentence stress as a grouping variable, results are less conclusive than was the case in the pooled data analysis. In some cases, sample size was too small to permit statistical comparison. Thus, a comprehensive assessment of place and manner effects across accented and unaccented tokens for all the vowels is not possible. Since a three-way stress comparison is prohibited for the same reason, results are reported for a two-way comparison between combined nuclear and non-nuclear accented tokens and unaccented tokens. Results of the analyses of variance are presented in Tables 6.15:16.

6.5.1 F2 as a function of adjacent consonantal place of articulation

In general, results are consistent with the expectation that unaccented vowel tokens show a greater shift in value as a function of context than the corresponding accented tokens. In some cases this also results in greater overall separability in midpoint value. The results also confirm that differences between accented and unaccented vowels with respect to the directionality of coarticulatory effects are largely differences in the magnitude of effects. In general, accented and unaccented tokens for a given vowel category show the same pattern with respect to the number and type of contexts that exert the greatest influence at both the carryover and anticipatory level. As in the pooled data, any differences in the relative magnitude of effects across accent conditions as indicated in the regression analyses, are largely accounted for by differences in the distribution of these contexts.

In the case of /i/, greater context-sensitivity for unaccented tokens predictably manifests itself in lower F2 midpoint values relative to the values for the corresponding accented tokens. At the carryover level, this also results in greater overall separability in F2 midpoint value for unaccented compared with accented tokens. The only pairwise difference to attain significance in the accented condition is for tokens in the context of a preceding alveolar compared with tokens in the context of a preceding apical, higher values being obtained in the alveolar context ($p < .05$). Within the unaccented condition, significantly higher values are obtained for tokens in both velar and alveolar contexts compared with tokens in palatal and dental contexts. This is chiefly due to lower values being obtained in the unaccented palatal and dental contexts. Values are also significantly lower for tokens following apicals than for tokens following alveolars. The lowering effect of a preceding /w/ is also greater for unaccented tokens than for accented tokens (mean 1900 Hz compared with 2084 Hz) resulting in significantly lower values for tokens in this context compared with tokens in the context of a preceding labial ($p < .001$) or apical ($p < .05$).

In a comparison of group means across accent conditions, the results confirm that labio-velar /w/ and apicals exert a stronger influence on unaccented /i/ tokens than on accented /i/ tokens. Values are significantly lower in both unaccented contexts although the significance level is higher in the case of apicals ($p < .01$) than in the case of /w/ ($p < .1$). It is likely that differences in the level of significance between comparisons reflect differences in the sample size between contexts (69 and 162 apical contexts compared with 10 and 19 labio-velar contexts for the accented and unaccented conditions respectively).

With respect to the influence of following context, there is no difference between accent conditions in terms of overall separability, none of the pairwise comparisons within either condition show significant differences. However, significantly lower values are obtained for unaccented tokens in alveolar contexts compared with the corresponding accented tokens ($p < .01$). As in the pooled data, values for tokens preceding apicals tend to be lower than for tokens in all other contexts. While the difference in value within either accent condition fails to attain significance, unaccented tokens in this context show significantly lower values than accented tokens preceding alveolars and dentals ($p < .05$) and accented tokens, significantly lower values than unaccented tokens in velar ($p < .001$), alveolar ($p < .01$) and labio-dental contexts ($p < .001$).

Unaccented /ɪ/ tokens also show greater differentiation between individual contexts than accented /ɪ/ tokens. At the carryover level, this reflects a relatively greater raising and lowering influence of velars and apicals respectively such that unaccented tokens show significantly higher values in the context of a preceding velar than in any other context and significantly lower values following apicals than in all contexts other than /w/. Accented versus unaccented comparisons, however, are only significant in the case of apicals. An accented versus unaccented comparison of labio-velar contexts is not possible due to there being only one example of accented /ɪ/ occurring in this context.

Both accented and unaccented /ɪ/ tokens show greater separability as a function of C^2 place of articulation than as a function of C^1 place of articulation. However, the lowering effect of a following labial ($p < .001$) and apical ($p < .05$) is significantly greater for unaccented compared with accented tokens. Lower mean values for the unaccented condition are also obtained for tokens in all the other contexts, most notably preceding dentals and labio-velars. However, the accented versus unaccented pairwise comparisons do not show a significant effect in these cases.

Unaccented /a/ tokens are also subject to a greater raising effect from a preceding velar than the corresponding accented tokens. Values for unaccented tokens are significantly higher following velars than in all other contexts other than a preceding palatal or alveolar context. Values are also lower for unaccented tokens following labio-dentals than following alveolars, velars and labials. Apart from a significant lowering effect of a preceding labio-dental, accented /a/ tokens show little separability in F2 midpoint value as a function of C^1 place of articulation. For both accented and unaccented /a/ tokens, the only significant difference in value as a function of C^2 place of articulation occurs for tokens preceding apicals. The amount of shift in F2 midpoint value is similar in both cases.

The raising influence of a preceding alveolar is stronger for unaccented /ʌ/ tokens than for accented /ʌ/ tokens. This is reflected in a significant difference in value both between unaccented tokens in this context and the corresponding accented tokens ($p < .05$) and between unaccented tokens in this compared with unaccented tokens in the labial ($p < .001$) and apical ($p < .01$) contexts. For accented tokens, there is no difference in value for tokens following alveolars compared with tokens following labials or apicals. While unaccented /ʌ/ tokens show higher mean values than the corresponding accented tokens for all C^2 contexts other than the apical context, none of the accented versus unaccented pairwise comparisons show significant differences.

A relatively stronger raising effect from an adjacent alveolar on unaccented compared with accented tokens is also evident in the case of /ɒ/. Unaccented /ɒ/ tokens in the context of a preceding alveolar show significantly higher values than accented tokens in the velar context ($p < .05$). Within the accented condition, there is no difference in value between tokens following alveolars and velars. Higher values are also obtained for unaccented tokens in the context of a following alveolar resulting in significant pairwise differences between unaccented tokens in alveolar contexts and accented tokens in dental ($p < .001$) and in velar and labial contexts ($p < .01$).

Although unaccented /ɔ/ tokens show higher mean values than the corresponding accented tokens there are no significant pairwise differences between accent conditions. With the exception of tokens in the labial context, unaccented /ʉ/ tokens show lower mean values as a function of C^1 place of articulation than the corresponding accented tokens. However, the difference in value is only significant for tokens in the alveolar and velar contexts. There is no significant difference in F2 value for accented or unaccented /ʉ/ tokens as a function of C^2 place of articulation.

In the case of /ɛ/, lower values are obtained for unaccented tokens in the context of a following dental than in all contexts other than the labial and labio-dental contexts, resulting in significant pairwise differences not evident in the accented condition. However, values are not significantly lower than the corresponding values for accented tokens in this context. In the case of /ʊ/ and /ɜ/, comparison of contexts across accent conditions is not possible due to small sample sizes.

6.5.2 F1 as a function of adjacent consonantal manner of articulation

The pairwise comparisons of group means for F1 are consistent with the results for F2 insofar as they show that accented and unaccented tokens differ largely with respect to amount of coarticulation rather than the number and type of contexts which exert a coarticulatory influence.

The raising influence of a preceding apical is stronger for unaccented /i/ tokens than for accented /i/ tokens ($p < .001$). There is also a tendency for unaccented /i/ tokens in the context of a preceding glide to be characterised by higher values than the corresponding accented tokens ($p < .1$). Relatively higher values for unaccented /i/ tokens in these contexts as well as in the fricative context result in significant pairwise differences between unaccented tokens in these compared with tokens in the stop context ($p < .001$). Within the accented condition, lower values only occur for tokens following stops compared with tokens following nasals ($p < .01$). At the anticipatory level, there is less distinction between accented and unaccented tokens. In general, higher values are obtained for tokens preceding nasals than for tokens in other contexts. Accented tokens in the context of a following nasal display higher values than accented tokens preceding stops and fricatives ($p < .05$) and higher values than unaccented tokens preceding stops ($p < .05$), fricatives ($p < .01$) and nasals ($p < .001$). Higher values are also obtained for unaccented tokens preceding nasals than for unaccented tokens preceding stops ($p < .01$) or for accented tokens preceding stops ($p < .05$).

In the case of /ɪ/, lower values occur for unaccented tokens compared with accented tokens in the context of a preceding fricative ($p < .01$). Otherwise there are no significant pairwise differences at the carryover level between accented and unaccented tokens in equivalent contexts. Relatively lower values for unaccented tokens in the fricative context results in a significant pairwise difference in value between these and unaccented tokens in the context of a preceding nasal ($p < .001$) or liquid ($p < .01$). In the accented condition, there is no significant difference in value for tokens in these contexts. Conversely, the relatively higher values for accented tokens following fricatives results in a significant pairwise difference between these and tokens following stops ($p < .05$). There is no difference in value between unaccented tokens in stop and fricative contexts. Unaccented tokens in the context of a preceding liquid and glide also show a significant difference in value ($p < .001$) which is not apparent in the accented condition although there is a significant difference in value between accented tokens following glides and unaccented tokens following liquids ($p < .01$) and between accented tokens following liquids and unaccented tokens following glides ($p < .001$). In both cases, tokens in the liquid contexts display higher values than tokens in the glide contexts.

With respect to the influence of following context, both accented and unaccented tokens display a high degree of separability in midpoint value. Within both accent conditions, higher values occur in the context of nasals and liquids than in the context of stops and fricatives. Values are higher for accented tokens in the context of a following nasal than for the corresponding unaccented tokens ($p < .01$). Otherwise there are no significant differences between accented and unaccented tokens in equivalent contexts.

Significantly lower values are obtained for unaccented compared with accented /ɛ/ tokens in the context of a preceding glide ($p < .01$). While there are no other significant pairwise differences between accented and unaccented tokens in equivalent contexts, generally higher F1 values for accented tokens result in significant differences between accented tokens in preceding stop contexts and unaccented tokens in preceding glide contexts ($p < .01$), between accented tokens in preceding nasal contexts and unaccented tokens in preceding stop and glide contexts ($p < .001$) and preceding fricative contexts ($p < .01$) and between accented tokens in preceding fricative contexts and unaccented tokens in preceding glide contexts ($p < .01$). Within each accent condition, however, there are no significant differences in F1 midpoint value as a function of C^1 manner of articulation.

At the anticipatory level, unaccented tokens show significantly lower values in the context of a following stop ($p < .05$) or nasal ($p < .001$) than the corresponding accented tokens. The values for unaccented tokens preceeding stops are also lower than the values for accented tokens preceding fricatives ($p < .01$) and liquids ($p < .001$). While accented tokens show slightly lower values preceding stops than preceding liquids

($p < .1$), there is no difference between accented tokens in the context of a following stop compared with a following fricative.

In the case of /a/, a significant main effect of sentence stress results in lower F1 values for unaccented tokens in preceding fricative contexts than for accented tokens in the context of a preceding stop ($p < .001$), fricative ($p < .01$), liquid ($p < .001$) or nasal ($p < .05$). Values are also significantly lower for unaccented tokens following stops than for accented tokens following liquids ($p < .01$). There are no significant pairwise differences within either the accented or unaccented condition. There is also a greater lowering effect for unaccented tokens from following stop contexts. Values for unaccented tokens in this context are significantly lower than the values for the corresponding accented tokens ($p < .001$) and lower also than the values for accented tokens in the context of a following nasal ($p < .001$) or fricative ($p < .01$).

F1 values are significantly lower for unaccented /ʌ/ tokens in the context of a preceding stop or fricative than either accented or unaccented /ʌ/ tokens in the context of a preceding nasal. Unaccented tokens following liquids also tend to show lower values than accented tokens in the nasal context ($p < .1$). Lower values for unaccented tokens also occur as a function of C^2 manner of articulation. Unaccented tokens show lower values in the context of a following stop ($p < .001$), fricative ($p < .01$), nasal ($p < .1$) or liquid ($p < .05$) than accented tokens in the context of a following nasal. While there is also a difference in value for accented tokens preceding nasals compared with accented tokens preceding stops ($p < .05$) and liquids ($p < .001$), there is no difference in value for accented tokens in following nasal compared with following fricative contexts.

Values are also significantly lower for unaccented /ɒ/ tokens in the context of a preceding nasal or stop ($p < .001$) or fricative, liquid or glide ($p < .05$) than for accented tokens in the context of a preceding nasal. Within the accented condition, only tokens in preceding stop and fricative contexts show significantly lower values than tokens in the nasal context ($p < .01$). There are no differences as a function of degree of sentence stress at the anticipatory level.

There is no variation in F1 for /ɑ/ tokens as a function of either sentence stress or context at either the carryover or anticipatory level. /ɔ/ tokens show a significant main effect of context at the carryover level but no stress effect while /ɜ/ tokens show a significant main effect of context but no stress effect at the anticipatory level. For /ʉ/, there is a significant difference in value for accented tokens in the context of a preceding stop compared with unaccented tokens following glides ($p < .01$), with higher values being displayed by the accented tokens. Lower values are also obtained for unaccented tokens following glides compared with unaccented tokens in the context of a preceding liquid ($p < .05$). There are no significant pairwise differences within the accented condition. There is no difference in value for accented compared with unaccented tokens as a function of following context.

6.6 Discussion

6.6.1 Duration

The results of the present study accord with the literature in showing generally longer durations and more peripheral formant values for sententially stressed compared with sententially unstressed vowels. In general, unaccented vowel tokens in the present data also show greater overall variability and greater context-dependency than their accented counterparts resulting in a lesser overall degree of acoustic contrast between vowels.

The observed durational difference between accented and unaccented vowels, in many cases, reflects a difference for the nuclear accented compared with the unaccented condition. The vowels /a, ɜ, ɑ, ʌ, ɔ/ show no difference in duration as a function of bearing a non-nuclear accent compared with no accent. /i/, /ɒ/ and /ʉ/ also show a relatively greater difference in duration between nuclear accented and unaccented tokens than between non-nuclear accented and unaccented tokens. The mid-high, lax vowels /ɪ/ and /ʊ/ are the only vowels to show comparable durations for tokens in the two accented conditions.

The lack of a significant difference in duration between non-nuclear accented and unaccented vowel tokens may reflect an averaging effect from other factors which influence vowel duration. Factors such as word structure, syllable structure, the location of clause and phrase boundaries and the manner and voicing characteristics of adjacent consonants (Klatt, 1976; Lehiste, 1960, 1972) are all known to affect vowel durations. In the present study, these variables are not controlled for and hence may serve to obscure stress-related durational differences. Given that sentence stress is a relational rather than an absolute property (Lehiste, 1970; Couper-Kuhlen, 1993), degree of stress and hence duration is also likely to vary across different utterances. The durational difference between nuclear accented and unaccented tokens is presumably not affected to the same extent because it is greater in absolute terms when all other factors are held constant.

Contrary to observations in the literature (Peterson & Lehiste, 1960; Klatt, 1973), in the present data, the relative degree of shortening as a function of reduced stress, is not uniform across all vowels. While the tense vowels /ɑ/, /ɜ/, /ɔ/ and the long, lax vowel /a/, show the greatest absolute shortening from the nuclear accented to the unaccented condition, it is the high, tense vowels /i/ and /u/ which show the greatest relative degree of shortening across the two accent conditions, followed by their lax counterparts /ɪ/ and /ʊ/. The long vowels /ɔ/, /ɜ/, /ɑ/, and /a/ show a slightly greater reduction in mean duration than /ɛ/ and /ʌ/ (between 1% and 3%) for unaccented tokens compared with nuclear accented tokens but a lesser decrease in duration than /ɒ/. The high and mid-high vowels together with the mid-low front vowel /ɛ/, also show the most difference in duration between non-nuclear accented and unaccented tokens. /ɑ/, /ɜ/, /ɔ/ and /a/ show the least difference in mean duration between non-nuclear accented and unaccented tokens due to relatively long durations being maintained for the latter.

The fact that the high and mid-high vowels display a relatively greater decrease in duration with reduced stress than the low vowels is consistent with the proposal that inherent segmental duration in conditioned by production constraints (Lindblom, 1967; Klatt, 1973). According to Lindblom, longer inherent durations for the low vowels reflect longer time requirements for jaw movement relative to tongue movement. Extrapolating from this, the relatively long inherent duration for the mid-low vowel /ɔ/ arguably reflects the cumulative time requirements of jaw movement coupled with lip rounding.

However, while it may be the case that "incompressibility is relative to the inherent duration of a phonetic segment and reflects a minimum time of execution of the required articulatory program." (Klatt, 1973, p. 1103), the present results also support the argument that the durational hierarchy is maintained on the basis of perceptual constraints, specifically, to preserve the distinction between the spectrally similar tense/lax pairs of vowels which are otherwise subject to a loss of acoustic contrast with a decrease in stress. This would also account for the relatively long durations for unaccented /ɜ/ tokens. The durational distinction between the tense/lax pairs /i/-/ɪ/ and / u/-/ʊ/ would appear to be maintained through a comparable degree of shortening for each pair member.

The pattern of results with respect to the amount of variance in vowel midpoint value that is explained uniquely by duration is similar for both accented and unaccented tokens. Duration accounts for a relatively larger proportion of the variance in F2 midpoint value for the peripheral vowels /i, ɑ, ɒ/ and /ɔ/ than for the more central vowels in both accent conditions. With respect to F1, the low and mid-low vowels also show a larger unique contribution of duration than the high and mid-high vowels in both accent conditions.

In the case of /i/, /ɑ/ and /ɒ/, the relative contribution to the F2 prediction made by duration is larger for accented than for unaccented tokens. Similarly, with the exception of /ɛ/, /ɜ/ and /ɑ/, the relative contribution of duration to the F1 prediction for vowels is also larger for accented compared with unaccented tokens. Assuming that the size of the durational effect is conditioned by the degree of compatibility between vowels and consonants (see section 5.5), this result is consistent with the view that accented vowels are characterised by more extreme articulatory configurations than unaccented vowels.

6.6.2 Overall variability

In terms of overall variability in formant frequency value, the pattern of reduction in variability with an increase in sentence stress is consistent with expectations. Nuclear accented tokens generally display less variability than either the corresponding non-nuclear accented or unaccented tokens. With the exception of /ɜ/, non-nuclear accented tokens are also less variable than their unaccented counterparts. The lesser degree of variability for non-nuclear accented tokens, particularly for those vowels which show no durational difference between the non-nuclear accented and the unaccented condition, further supports the suggestion that durational differences between non-nuclear accented and unaccented tokens may be obscured through the influence of other factors. However, it may also reflect the fact that degree of context sensitivity can vary independently of vowel duration.

As with duration, the degree of reduction in overall variability as a function of increased stress is variable across vowels. In this case, the long vowels /ɔ/, /ɑ/ and /ɜ/ show a relatively greater degree of reduction as a function of stress than /i/ and /u/. Generally greater variation in the rank ordering of individual vowels across accent conditions with respect to overall variability than observed in the case of duration, may be attributed, in part, to differences in the frequency distribution of contexts between accent conditions. This may also account for differences in the relative amount of F1 compared with F2 variability for some vowels. Despite such differences, however, with the exception of /u/, the distinction between tense and lax vowel sets is maintained for both accented and unaccented tokens. Within each accent condition, the tense vowels as a set (including the long, lax vowel /a/) display lower overall variability than the lax vowels.

The diminished degree of acoustic contrast that accompanies a decrease in stress for both tense and lax vowel sets in the present data, conflicts with Fourakis' (1991) findings. While Fourakis reports a shrinkage in the size of the vowel space for lexically unstressed compared with lexically stressed vowel tokens, he found that individual vowels remained distinct. Given that Fourakis examined vowels in only two consonantal frames, the difference in results may be attributed to the greater range of contexts used in the present study and may, therefore, be taken as evidence in support of the view that the effects of stress and context are additive, that is, that unstressed vowels are more susceptible to coarticulatory effects than their stressed counterparts.

6.6.3 Degree of context-dependency

The results of the regression analyses and pairwise comparison of means tests confirm the greater context-sensitivity of unaccented vowels in the present data compared with accented vowels. With the exception of /ɛ/, /ɜ/, /ɑ/ and /ɔ/, unaccented vowel tokens show greater context-dependency along both F1 and F2 than the corresponding accented tokens. In general, nuclear and non-nuclear accented tokens pattern together. For those vowels which show a significant increase in context dependency with a decrease in sentence stress, the only vowels which do not show a significant difference in R^2 value between non-nuclear accented and unaccented tokens as well as between nuclear accented and unaccented tokens are /ɪ/, /ʊ/ and /ʌ/ in the case of F1 and /a/ in the case of F2.

Vowels also differ in terms of the amount of difference in degree of context dependency they show as a function of stress. For F2, the pattern of results within and across accent conditions suggests that vowels may be divided into three broad categories with respect to the question of inherent stability. Firstly, there are those vowels, namely /i/, /a/, /ɑ/

and /ɔ/ which show relatively low context-dependency in the pooled data analysis. The fact that /i/, /ɑ/ and /ɔ/ show comparatively little or no significant increase in degree of context-dependency for unaccented relative to accented tokens may be considered further evidence of their inherent robustness. While the long, lax vowel /a/ does show a comparatively large absolute difference in R^2 value across accent conditions, its relative position in the rank ordering from most to least context-dependent does not change appreciably for unaccented compared with accented tokens.

One possible explanation for the increase in R^2 value with an increase in stress observed for /ɔ/ is that it reflects a greater vowel-to-consonant effect than consonant-to-vowel effect for the accented relative to the unaccented tokens. Recasens (1991) reports less variation in consonantal values as a function of vowel-to-consonant coarticulation for those consonants which exert the strongest consonant-to-vowel coarticulatory effects. Since the regression analysis provides a measure of the strength of the relationship between onset/offset and midpoint value, a similar inverse relationship of reciprocal influences in the case of vowels, such that the vowels which show the least variation as a function of adjacent consonantal influence also exert the greatest vowel-to-consonant coarticulatory effects, would also manifest itself in relatively high R^2 values.

Accepting that stressed vowels exert a stronger coarticulatory influence than unstressed vowels (Fowler, 1981), the increase in R^2 value for accented /ɔ/ tokens relative to unaccented /ɔ/ tokens might be interpreted as reflecting a relatively stronger vowel-to-consonant effect for accented tokens rather than a stronger consonant-to-vowel effect for unaccented tokens. There is, however, a problem with this account in that the difference in degree of context-dependency between accented and unaccented /ɔ/ tokens largely reflects a difference between non-nuclear accented and unaccented tokens. Assuming that degree of coarticulatory influence increases with degree of stress, nuclear accented tokens would be expected to show the strongest vowel-to-consonant effect.

An alternative possibility is that the higher R^2 value for non-nuclear accented /ɔ/ tokens reflects differences in the frequency distribution of contexts across accent conditions. Non-nuclear accented /ɔ/ tokens show a relatively greater proportion of vowel contexts than either nuclear accented or unaccented /ɔ/ tokens (see Figure A 1.i). The majority of these are front vowels which exert a strong raising influence on the F2 values for /ɔ/.

Despite the relatively high degree of F2 context-dependency observed for non-nuclear accented /ɔ/ tokens relative to the corresponding nuclear accented and unaccented tokens, /ɔ/ continues to number among the four most robust vowels (i.e. /ɑ, ɒ, i, ɔ/) in the non-nuclear accented condition (see Figure 6.14).

The second posited category of vowels includes those vowels which represent the opposite extreme in terms of vowel robustness, namely /ɪ/, /ʊ/ and, to a lesser extent, /ɛ/. After schwa, /ʊ/ shows the greatest context-dependency along F2 in the pooled data analysis followed, in descending order, by /ɪ/, /u/ and /ɛ/. While /u/ exhibits higher overall F2 context-dependency than /ɛ/, its relative position in the hierarchy of least to most context-dependent shows a large difference between accented and unaccented conditions. In contrast, /ɛ/ shows no difference in level of context-dependency between accented and unaccented tokens. /ɪ/ and /ʊ/ similarly show a relatively small, although significant, stress effect compared with the stress effects observed for other vowels. Because /ɛ/ patterns with /ɪ/ and /ʊ/ in showing relatively high context-dependency for accented as well as for unaccented tokens, there is a sense in which it is generally less robust than /u/.

The third and final category encompasses the remaining vowels /u, ɒ, ɜ/ and /ʌ/ which show varying degrees of context-sensitivity and which, with the exception of /ɜ/, show a strong stress effect. Of these, the rounded vowels /u/ and /ɒ/ show the greatest difference in R^2 value as a function of stress. In each case, the difference in amount of context-dependency is in the expected direction and is of such a magnitude as to alter the relative positions of these vowels in the ranking from least to most context-dependent.

The long, central vowel /ɜ/ is also included in this category owing to the intermediate position it occupies in the rank ordering of vowels. Given its long inherent duration and comparatively low contextual variability, it cannot be included with the weakly specified vowels /ɪ, ɛ/ and /ʊ/. Relatively high levels of context-dependency in comparison with /ɑ, ɔ, a, i/, however, preclude it from inclusion in the robust category. It is, however, noteworthy, that while /ɜ/ shows a higher R^2 value in the accented condition compared with /ʌ/, /ɒ/ and /u/, in the unaccented condition, /ɜ/ displays a lower R^2 value than either /u/ or /ʌ/ and a similar value to /ɒ/.

To summarise, the differences in the ranking of vowels from least to most context-dependent as a function of sentence stress largely concern the lax vowels /ʌ/ and /ɒ/ and the tense vowel /u/. The lax, back vowel /ʌ/ occupies a higher position than the long, central vowel /ɜ/ in the accented condition but shows greater context-dependency than /ɜ/ in the unaccented condition. The lax, back vowel /ɒ/ shows less context-dependency than /i/ in the accented condition but greater context-dependency than /i/ in the unaccented condition. Accented /u/ tokens show a slightly lower level of context-dependency than accented /ɛ/ and /ɜ/ tokens. In the unaccented condition, however, /u/ shows a considerably higher level of context-dependency than these vowels.

The pattern of results for F1 shows some similarity to that observed for F2 but also some important differences. As in the case of F2, the long, back vowels /ɑ/ and /ɔ/ show the least context dependency for both accented (combined nuclear and non-nuclear) and unaccented tokens while the mid-high, lax vowel /ɪ/ shows the greatest context-dependency in both accent conditions. /ɑ/ and /ɔ/ also show no significant increase in degree of context-dependency with a decrease in stress. The long, front vowel /a/, however, generally occupies a lower position in the ranking from least to most context-dependent for F1 compared with F2. There is also a greater difference in the relative position it occupies in the hierarchy for accented compared with unaccented tokens in the case of F1. The relatively greater influence of stress on F1 compared with F2 in this case may reflect a greater duration dependence for jaw movement relative to tongue-body movement.

Large changes in rank position across accent conditions are also evident in the case of /u/, /ɒ/ and /ʊ/. For /u/ and /ɒ/, the change is in the expected direction with unaccented tokens showing greater context-dependency than accented tokens although for /u/ the difference largely reflects a difference between nuclear accented and unaccented tokens. In the case of /ʊ/, accented tokens display higher R^2 values than the corresponding unaccented tokens. This is true for tokens in both accented conditions although the difference is particularly marked for non-nuclear accented compared with unaccented tokens.

As in the case of the F2 results for /ɔ/, the increase in R^2 value for accented /ʊ/ tokens relative to unaccented /ʊ/ tokens might be construed as reflecting greater vowel-to-consonant coarticulation rather than greater consonant-to-vowel coarticulation. In addition to a decrease in the variability in midpoint value, nuclear and non-nuclear accented /ʊ/ tokens also show a decrease in the variability in onset and offset value relative to the range in onset and offset value apparent for the corresponding unaccented tokens. However, in this case, it is also more likely that the differences in R^2 value between accented and unaccented tokens simply reflect differences in the distribution of contexts across accent conditions. As Figure 6.25 demonstrates, non-nuclear accented tokens show a higher proportion of apical contexts which tend to raise F1 than either nuclear accented or unaccented tokens. Accented tokens also show a higher proportion of labial and velar stop contexts than the unaccented tokens.

The tense vowel /u/ and the lax vowels /ʌ/ and /ɛ/ show the highest level of F1 context-dependency after /ɪ/ in the pooled data analysis. Despite a higher absolute R^2 value for /u/ in the unaccented condition, /ʌ/ is arguably less robust than /u/ on account of the fact that it displays a relatively high R^2 value for accented as well as for unaccented tokens. Similarly, while the tense vowels /i/ and /u/ and the lax vowels /a/ and /ɒ/ all show greater context-dependency than /ɛ/ in the unaccented condition, the fact that /ɛ/ does not show a significant stress effect and occupies a relatively low position in the hierarchy from least to most context-dependent for accented tokens, suggests that it is also generally less robust than these vowels. While /ɪ/ shows a greater stress effect than /ʌ/ and /ɛ/, it also displays higher R^2 values for both accented and unaccented tokens.

In sum, the main differences between F1 and F2 in the ranking of vowels from least to most context-dependent concern the relative positions occupied by /ʊ/ and /a/. While /a/ displays greater absolute context-dependency along F2 than along F1, it occupies a lower relative position in the hierarchy of least to most context-dependent vowels for F1 than for F2, for both accented and unaccented tokens. Conversely, while /ʊ/ displays the highest R^2 value for F2 within each accent condition, unaccented /ʊ/ tokens show relatively little context-dependency along F1. Owing to the high ratio of unaccented to accented tokens for this vowel, /ʊ/ also displays a relatively low R^2 value for F1 overall.

In general, vowels show a relatively greater increase in amount of context dependency with a decrease in stress along F1 than along F2. In the case of /ɪ/ and /ʊ/, the stress effect is chiefly manifest in reduced context-dependency along F1. /ɪ/ and /ʊ/ show a decrease in degree of F1 context-dependency of 24% and 33% respectively for accented relative to unaccented tokens compared with a decrease in degree of F2 dependency of 1% and 4%. Given that increased stress is associated with increased vocal effort, a relatively greater stress effect for F1 than for F2 may be attributed to the fact that most of the energy of a vowel is contained within the first formant (Lehiste, 1970). A narrower specification for accented vowels along F1 than along F2 may also reflect the greater perceptual salience of F1 compared with F2.

6.6.4 Lexical distribution of vowels

The lack of uniformity in the relative size of the stress effect across vowels may reflect differences in their lexical distribution, that is, in their relative frequency of occurrence in function words compared with content words. Function words are likely to be subject to a greater degree of reduction than content words due to their relatively low informational content and their high frequency of occurrence (Bolinger, 1975). Thus, the relatively large increase in degree of context-dependency for unaccented relative to accented /u/ tokens may to some extent reflect the occurrence of /u/ tokens in function words such as "to", "into", "you", "who". /ɒ/ and /a/ also feature in high frequency function words: "of", "a", "at", "and". The full vowel percept in these cases may reflect phonemic restoration on the part of the labeller or the difficulty in discriminating between full and reduced vowels (see Van Bergem, 1995). However, owing to the lack of isomorphy between production and perception, vowels which qualify as full perceptually may show variable degrees of reduction at the production level.

The high degree of context-dependency observed for /ɪ/ and /ʊ/ might also be attributed to their relatively high occurrence in function words and suffixes. However, both vowels show a comparable amount of F2 context-dependency for accented as for unaccented tokens which suggests that they are generally more weakly specified along F2 than other vowels.

6.7 Summary

In this chapter, the role of sentence stress and the interaction of stress and context effects in conditioning vowel quality are examined. In line with predictions, unaccented vowel tokens generally show shorter durations, greater variability and, in the majority of cases, greater context-dependency than their accented counterparts.

Despite individual differences between vowels with regard to the amount of difference in degree of F2 context-dependency they show across accent conditions, the same general hierarchy of robustness is evident for both accented and unaccented tokens. The corner vowels /ɑ, i, a/ and the back, rounded vowels /ɒ/ and /ɔ/ show the least context-dependency in both accent conditions while the inherently short vowels /ɪ/, /ʊ/ and /ɛ/, together with /u/ in the unaccented condition, display the highest levels of context-dependency. The ranking of vowels in terms of degree of F1 context-dependency is more variable across accent conditions. The differences largely concern the vowels /ɒ/, /a/, /u/ and /ʊ/. For both accented and unaccented tokens, the back vowels /ɑ/ and /ɔ/ show the least F1 context-dependency while /ɪ/ together with /ʊ/ in the accented condition and /u/ in the unaccented condition, shows the highest F1 context-dependency. In general, vowels show lower context-dependency along F1 than along F2. However, they also tend to show a relatively greater stress effect along F1 than F2. The role of stress in conditioning vowel quality is further considered in Chapter 7.

Figure 6.2: Mean durations (msec) for nuclear and non-nuclear accented and unaccented vowel tokens

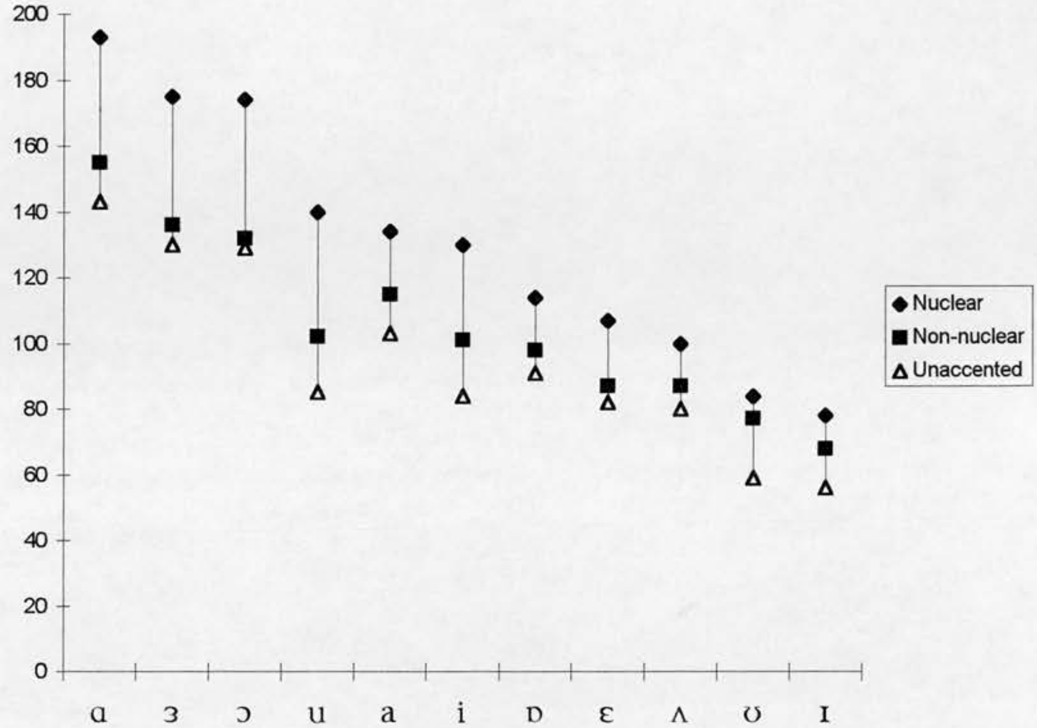


Figure 6.3: Percent reduction in duration between nuclear (Nuc), Non-nuclear (non-nuc) and unaccented (Un) accent conditions. This was calculated by dividing the difference in mean duration between, for example, the nuclear accented and the unaccented condition by the larger mean and multiplying the result by 100.

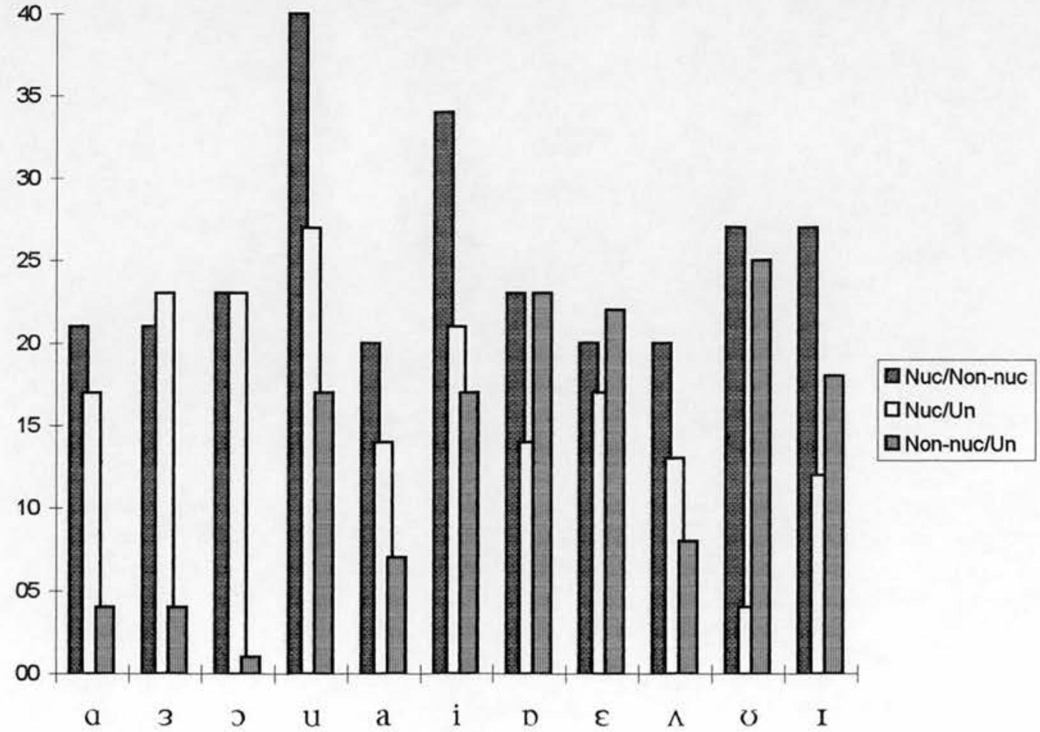


Figure 6.4: Tense vowel ellipses: combined nuclear and non-nuclear accented tokens.

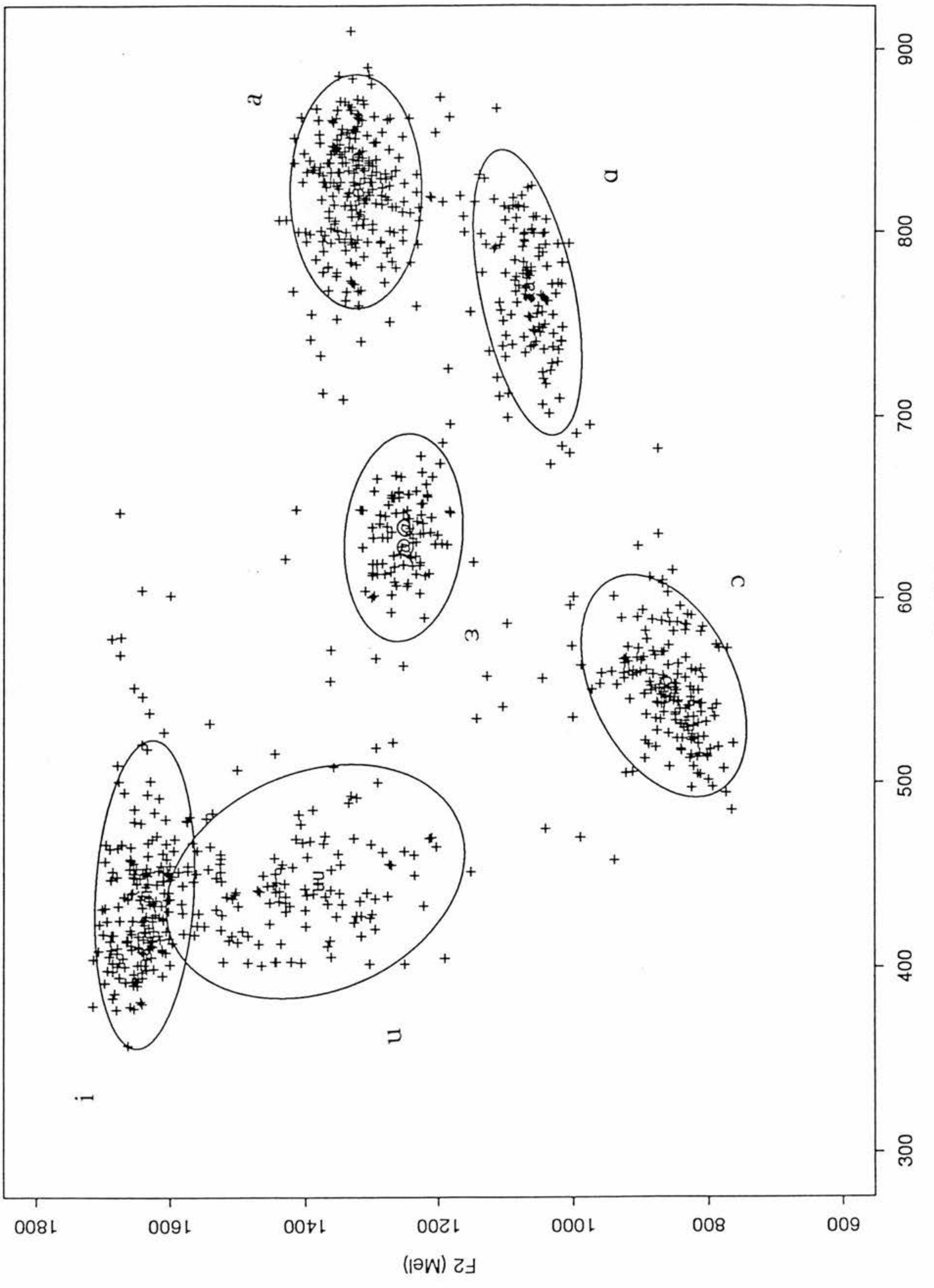


Figure 6.5: Tense vowel ellipses: unaccented tokens.

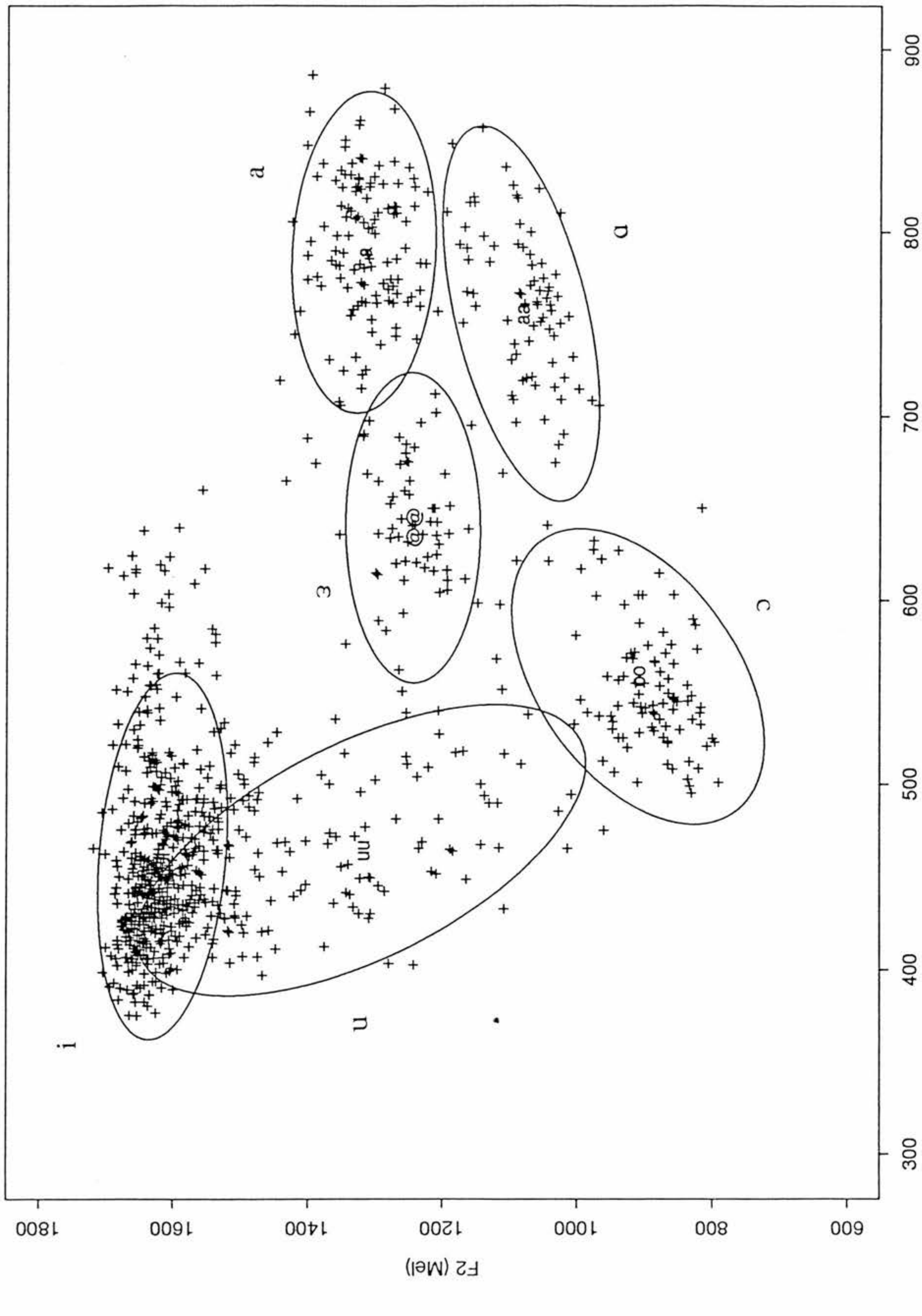


Figure 6.6: Lax vowel ellipses: combined nuclear accented and non-nuclear accented tokens.

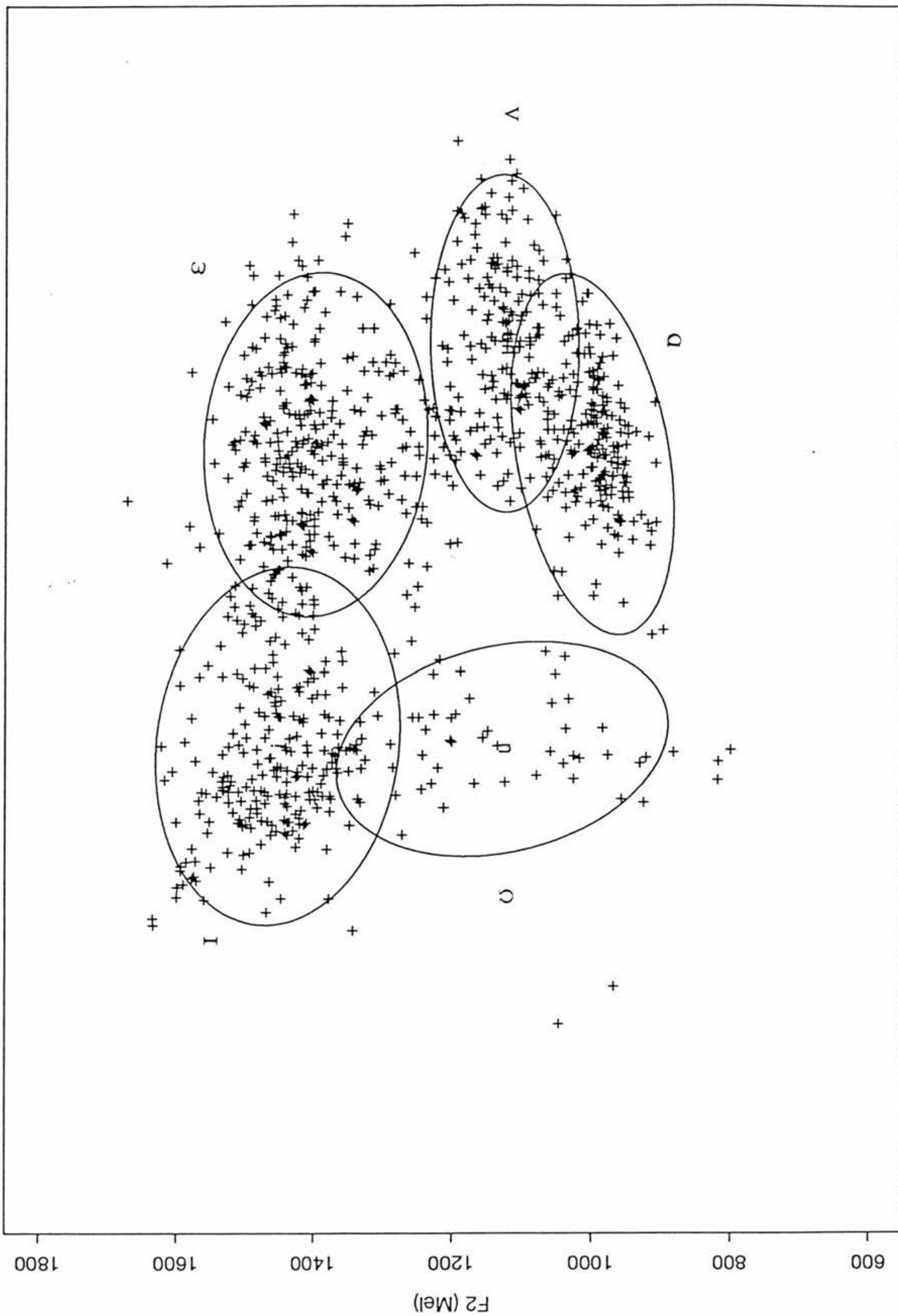


Figure 6.7: Lax vowel ellipses: unaccented tokens.

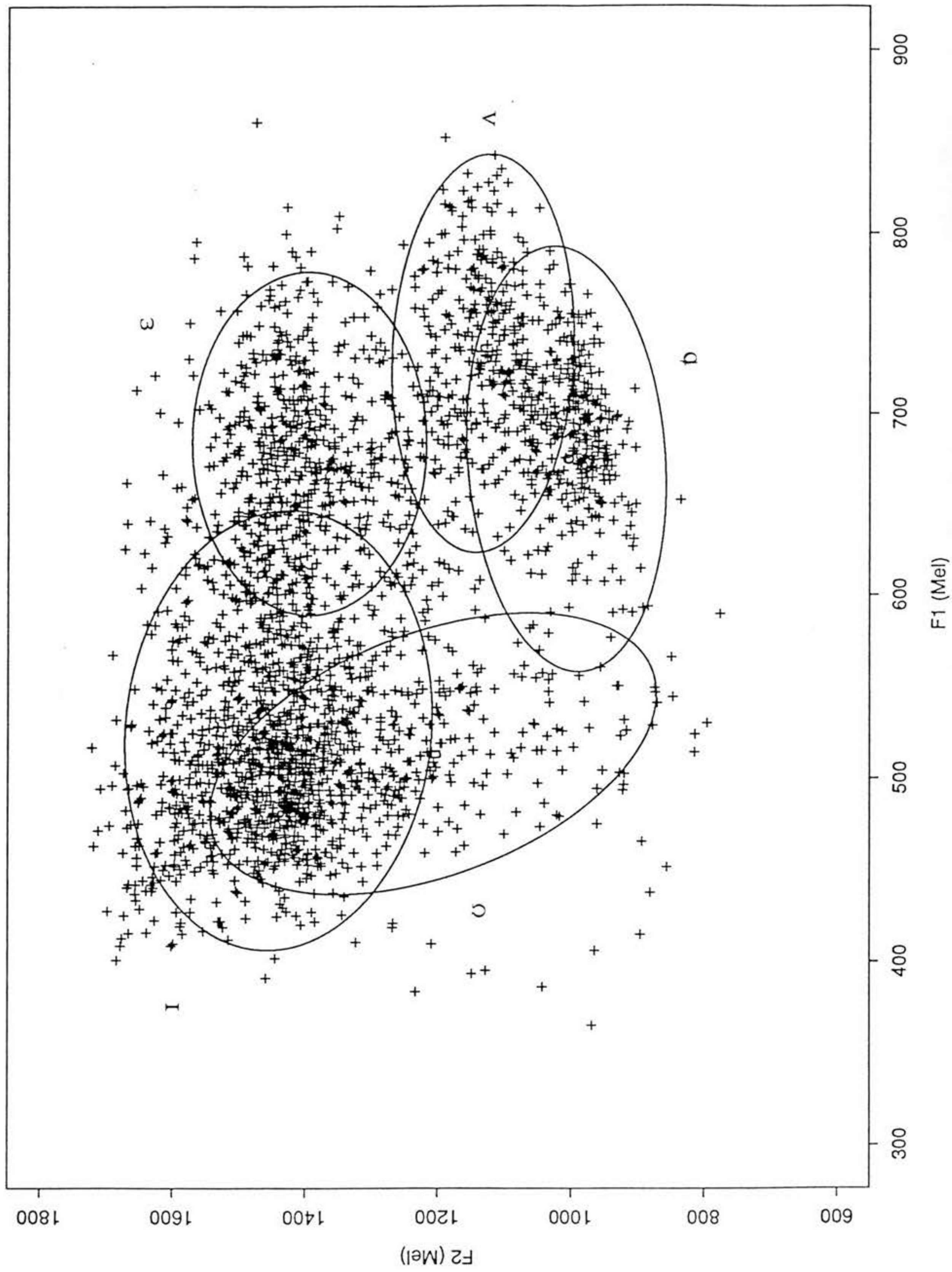


Figure 6.8: Tense vowel ellipses: nuclear accented tokens.

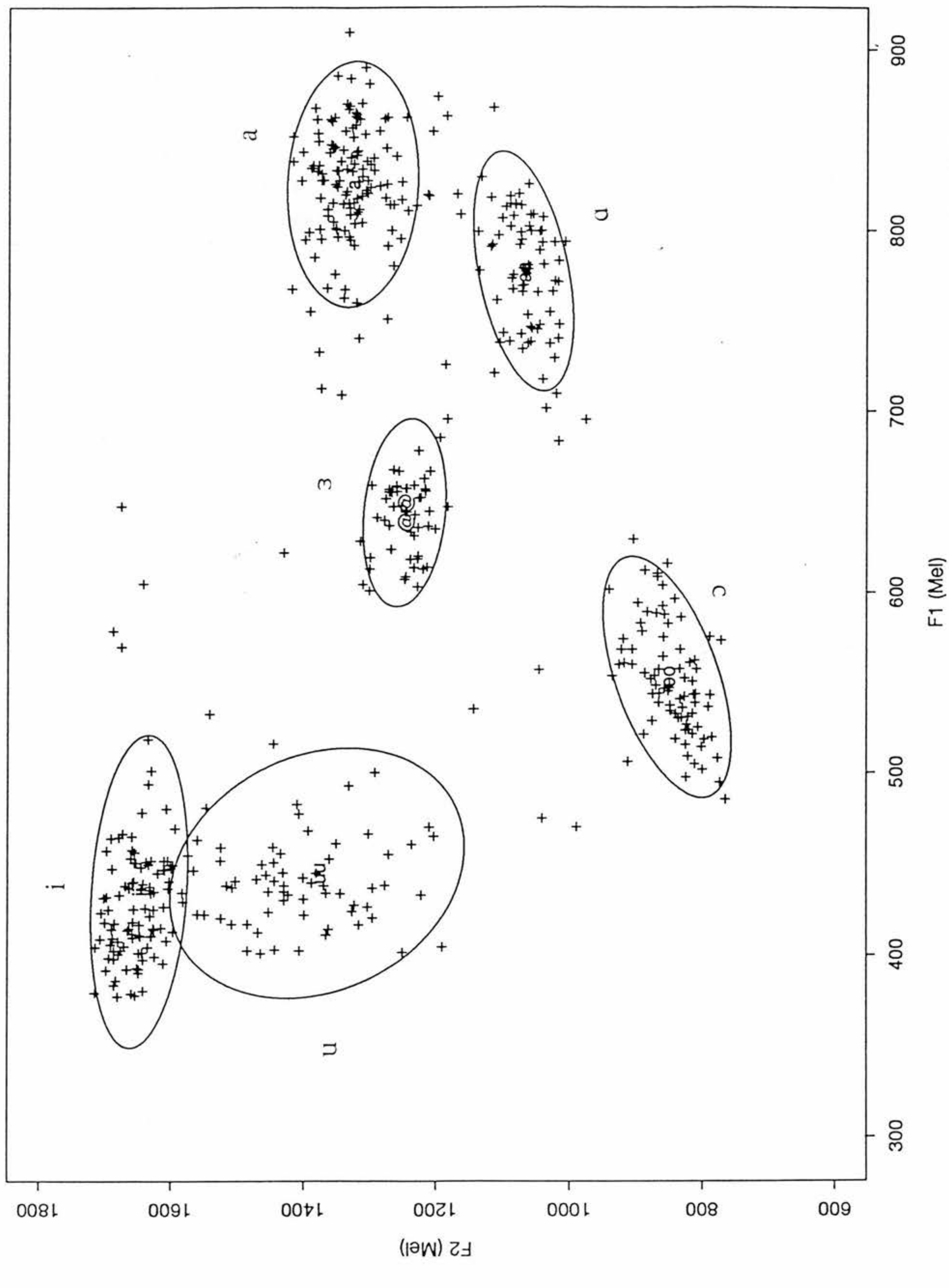


Figure 6.9: Tense vowel ellipses: non-nuclear accented tokens.

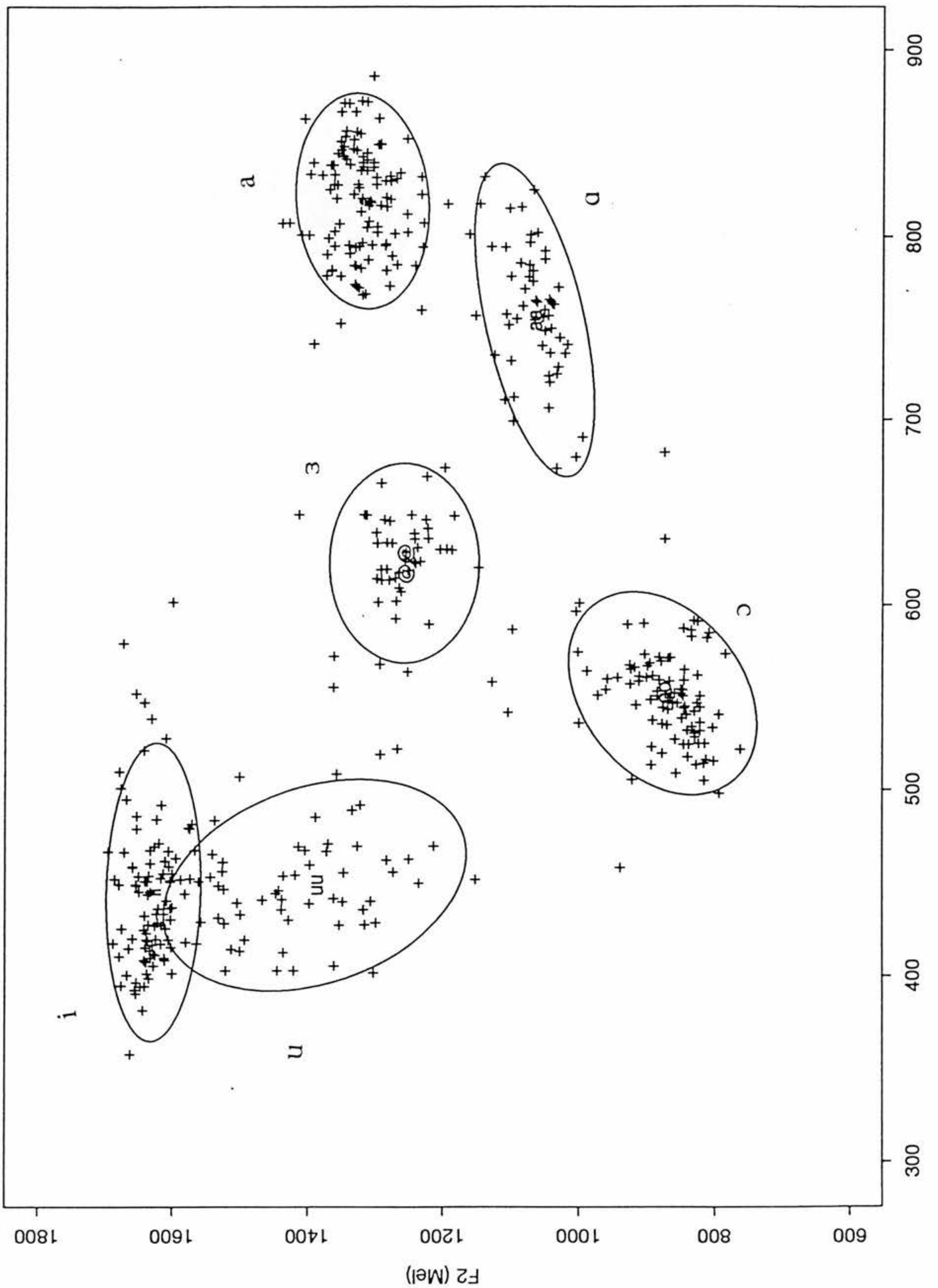


Figure 6.10: Lax vowel ellipses: nuclear accented tokens.

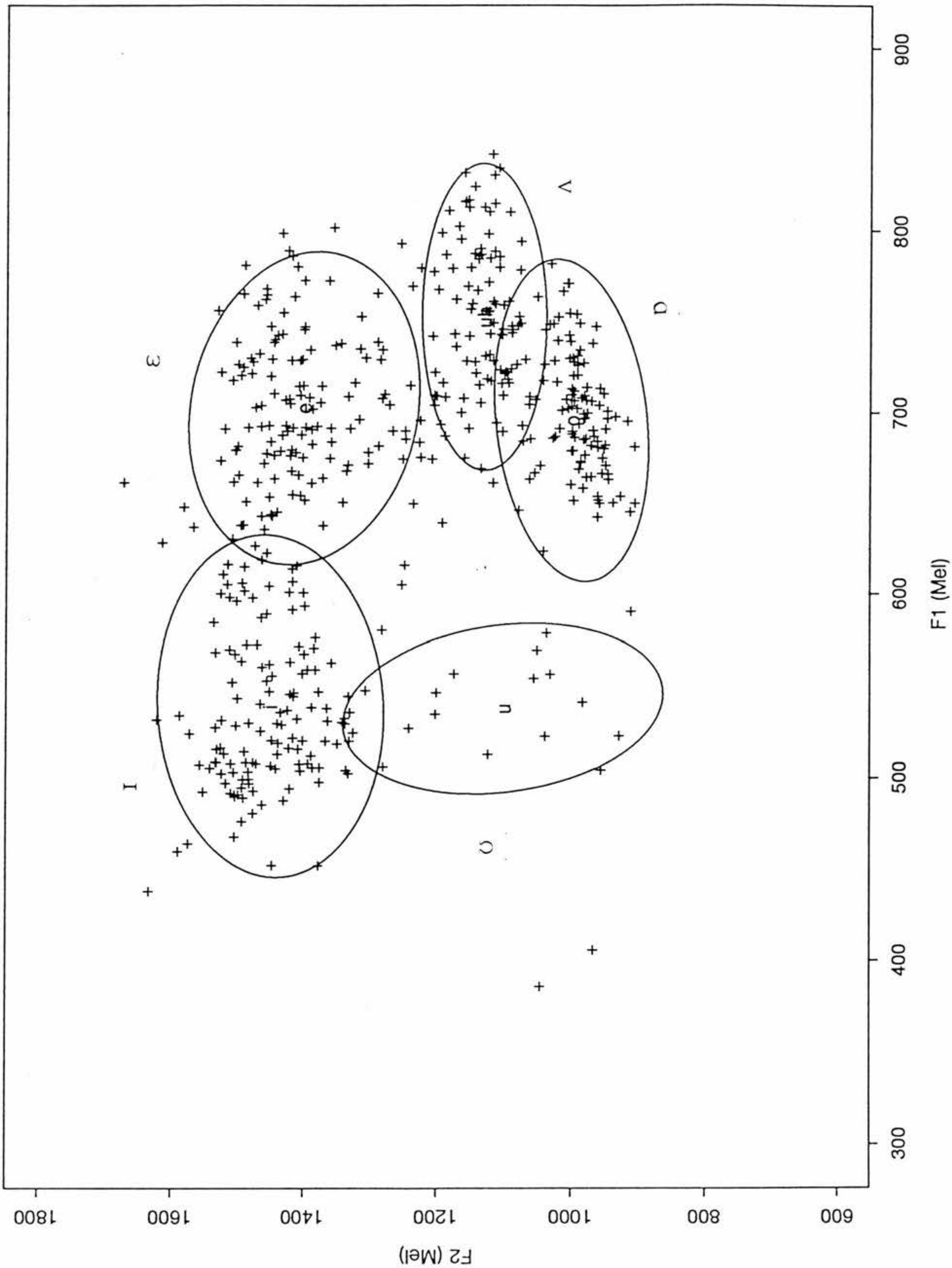


Figure 6.11: Lax vowel ellipses: non-nuclear accented tokens.

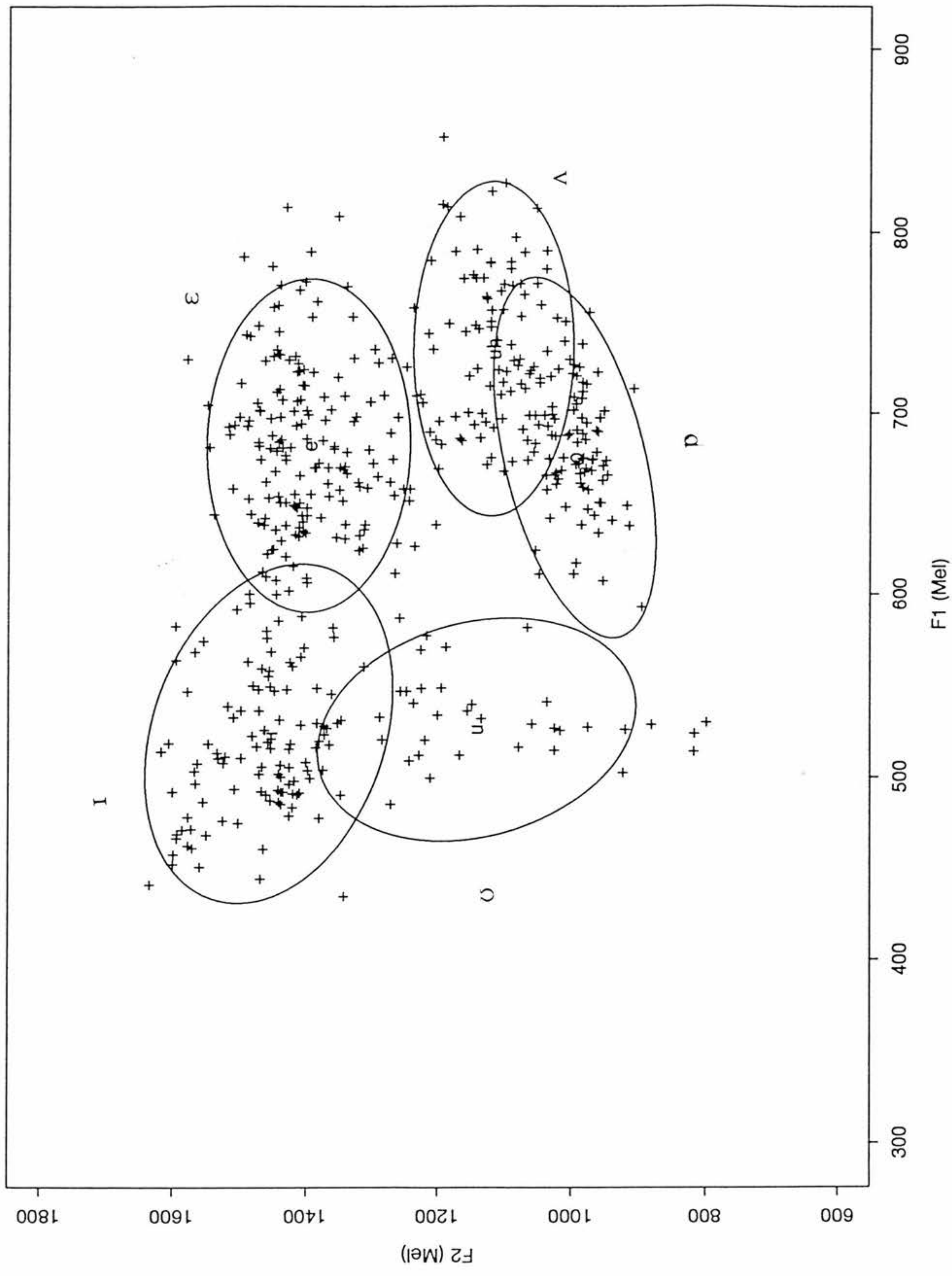
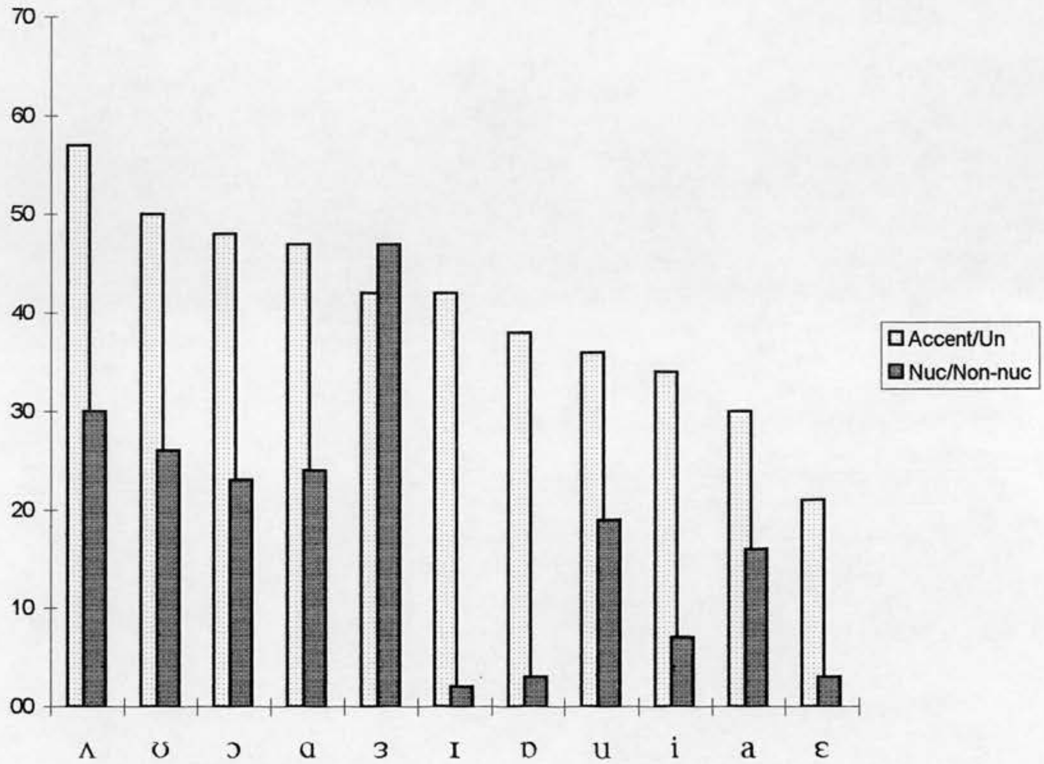


Figure 6.12: Percent reduction in overall variability as a function of stress

6.12a: combined nuclear and non-nuclear accented (Accent) vs unaccented tokens and nuclear (Nuc) vs Non-nuclear accented (Non-nuc) tokens



6.12b: Nuclear accented (Nuc) vs unaccented (Un) tokens and non-nuclear accented (Non-nuc) vs unaccented (Un) tokens

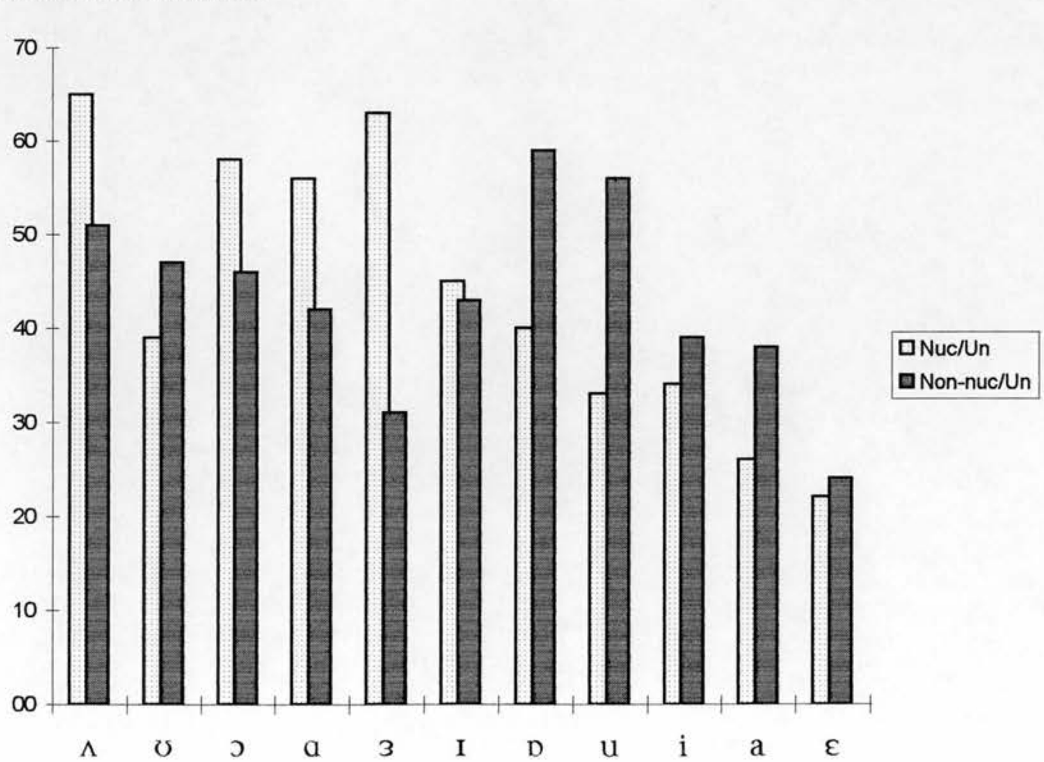
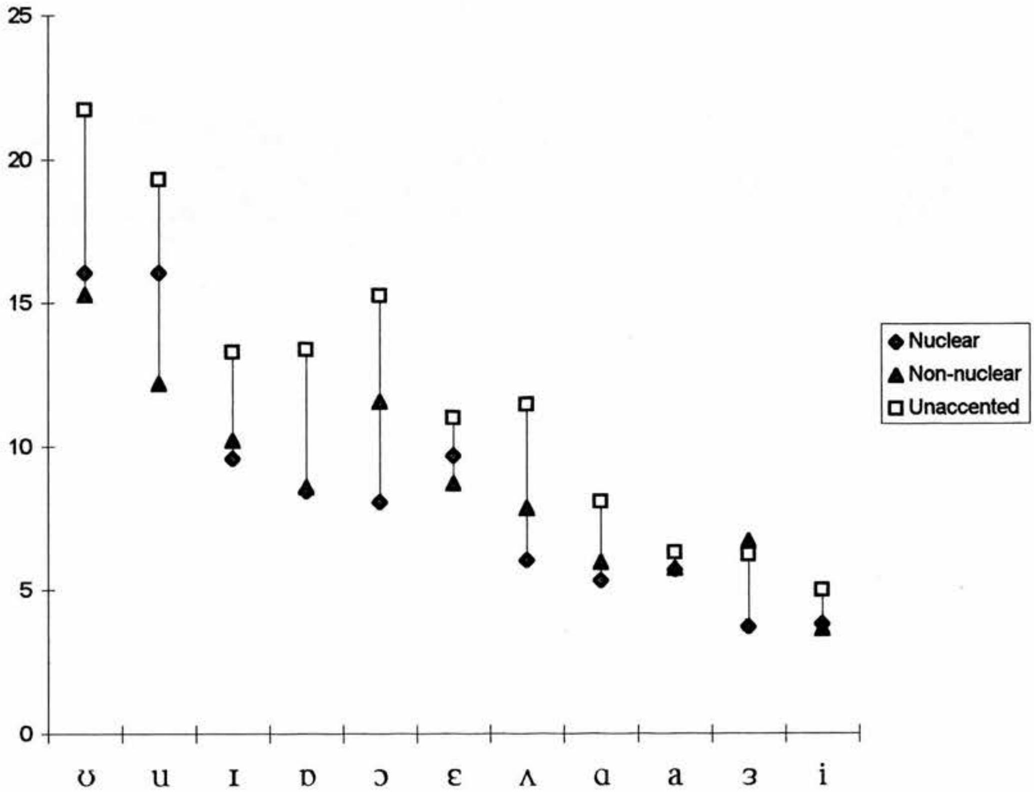


Figure 6.13: Coefficient of variance values for nuclear and non-nuclear accented and unaccented vowel tokens

6.13a: F2



6.13b: F1

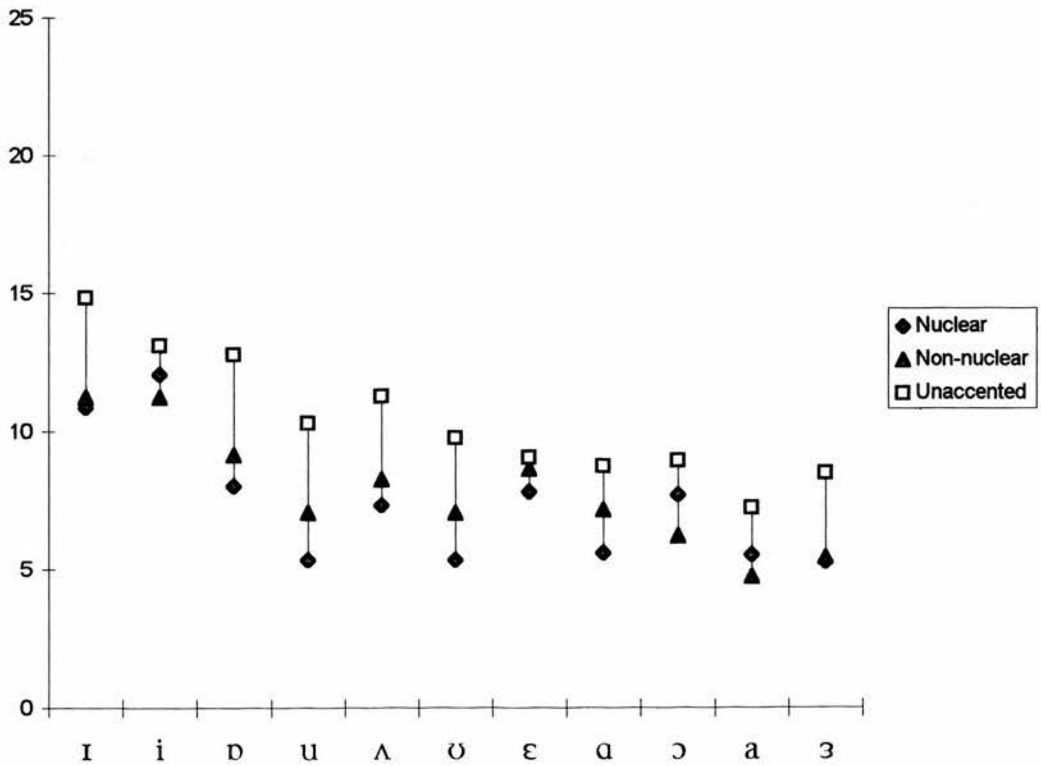
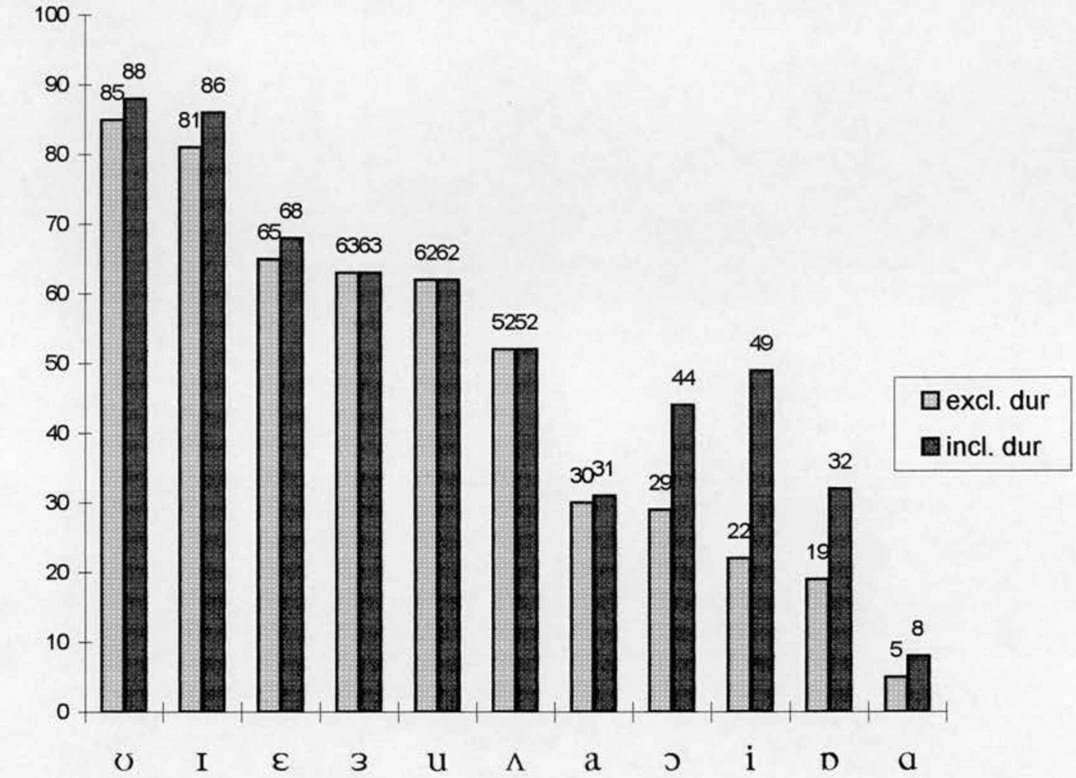


Figure 6.14: Proportion of explained variance: combined nuclear and non-nuclear accented tokens. The proportion of explained variance is given both including and excluding duration from the regression equation.

6.14a: F2



6.14b: F1

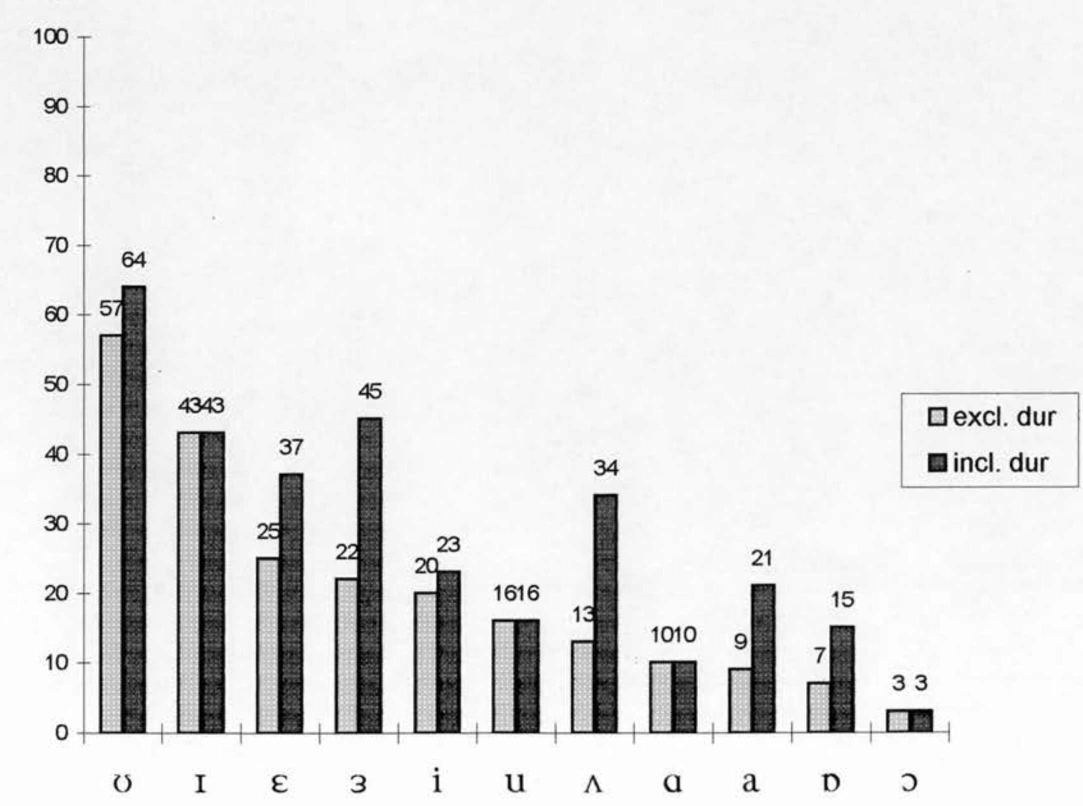
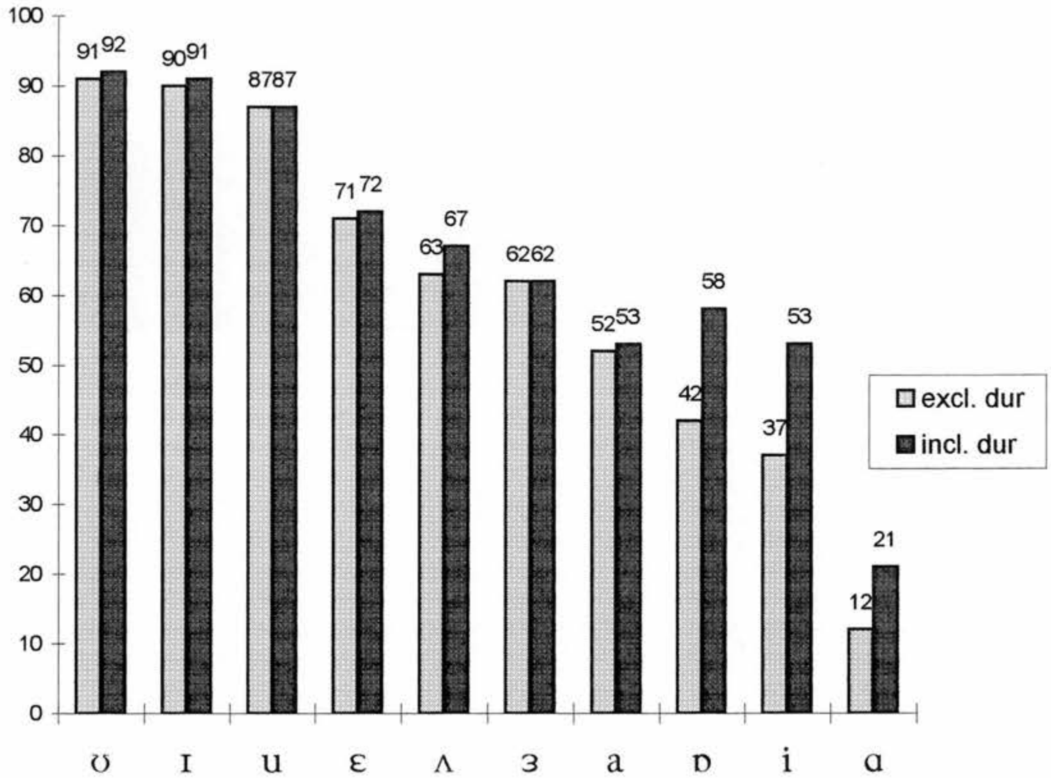


Figure 6.15: Proportion of explained variance: unaccented tokens. The proportion of explained variance is given both including and excluding duration from the regression equation.

6.15a: F2



6.15b: F1

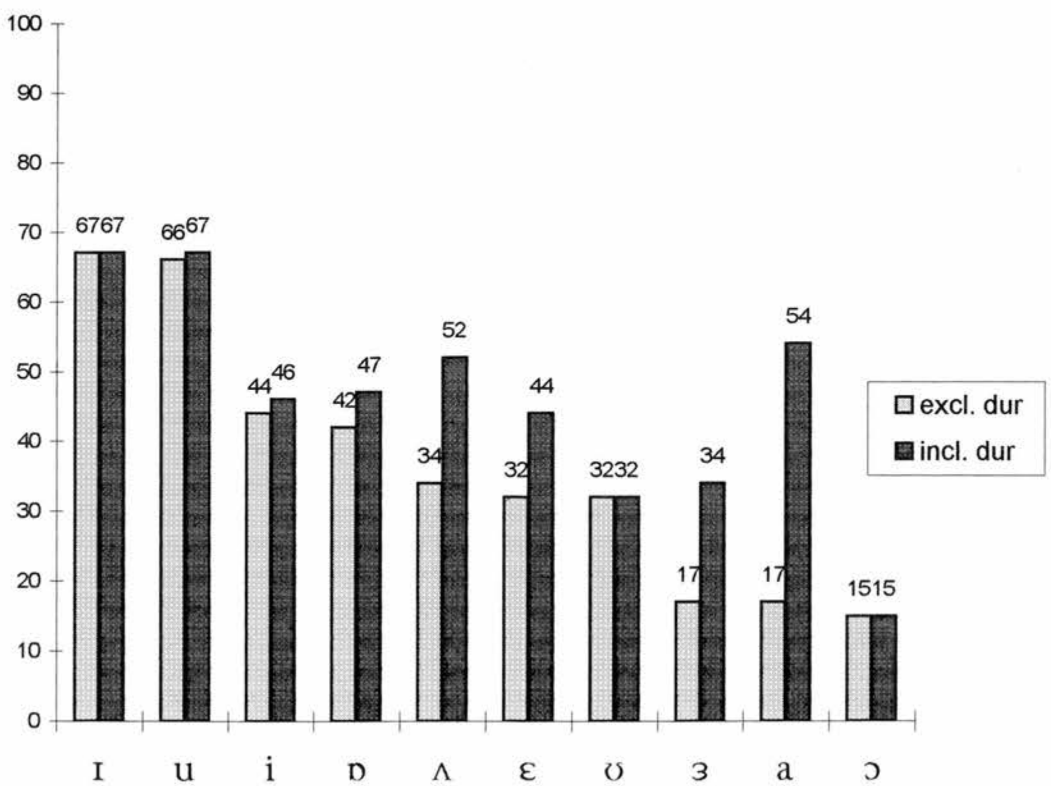
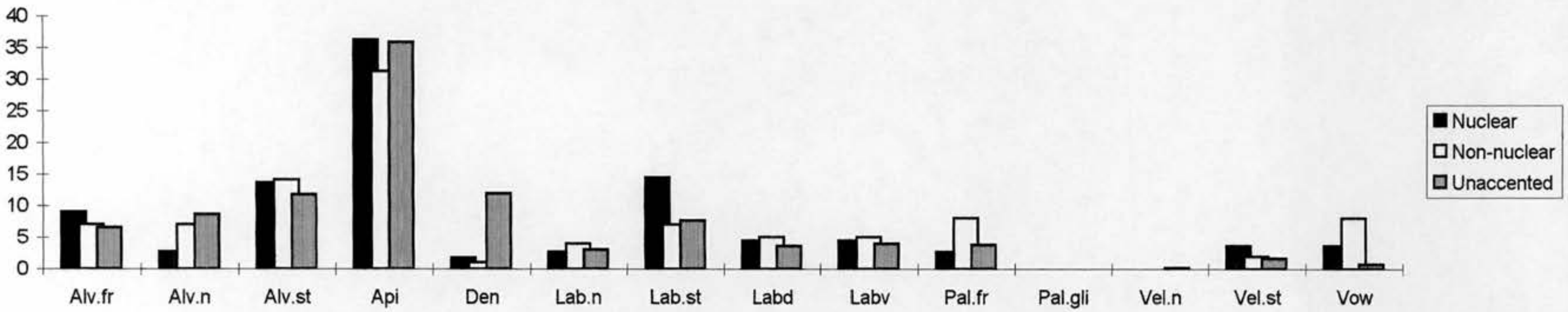


Figure 6.16: Proportional distribution of preceding and following contexts for nuclear and non-nuclear accented and unaccented /i/ tokens

6.16a: preceding contexts



6.16b: following contexts

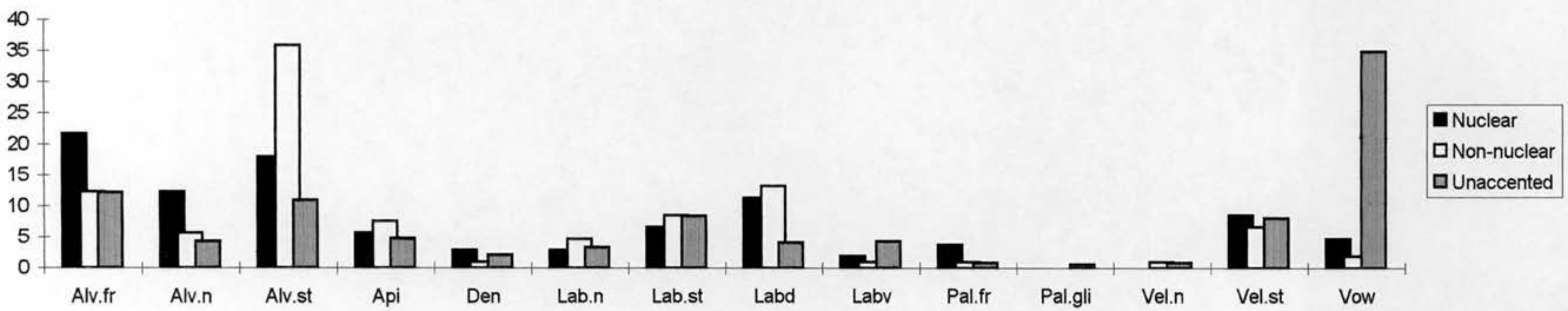
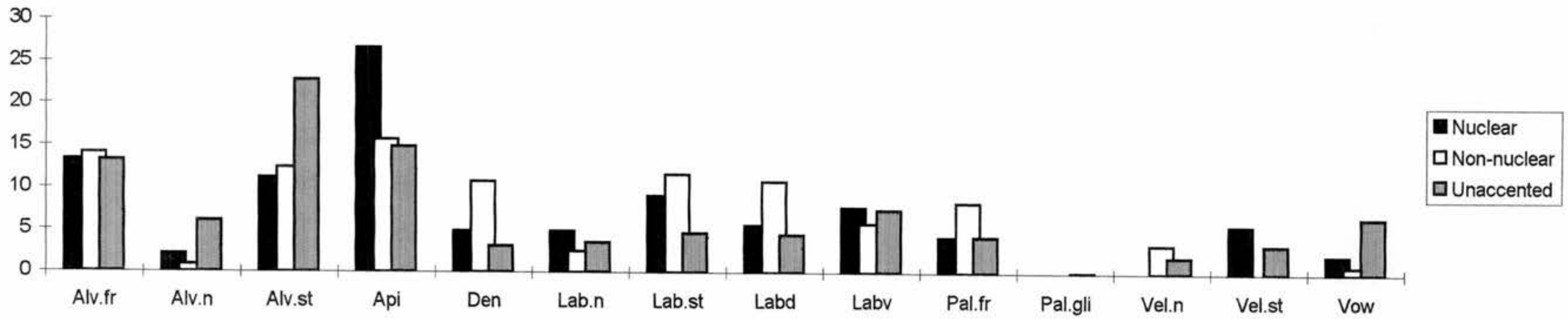


Figure 6.17: Proportional distribution of preceding and following contexts for nuclear and non-nuclear accented and unaccented /ɪ/ tokens

6.17a: preceding contexts



6.17b: following contexts

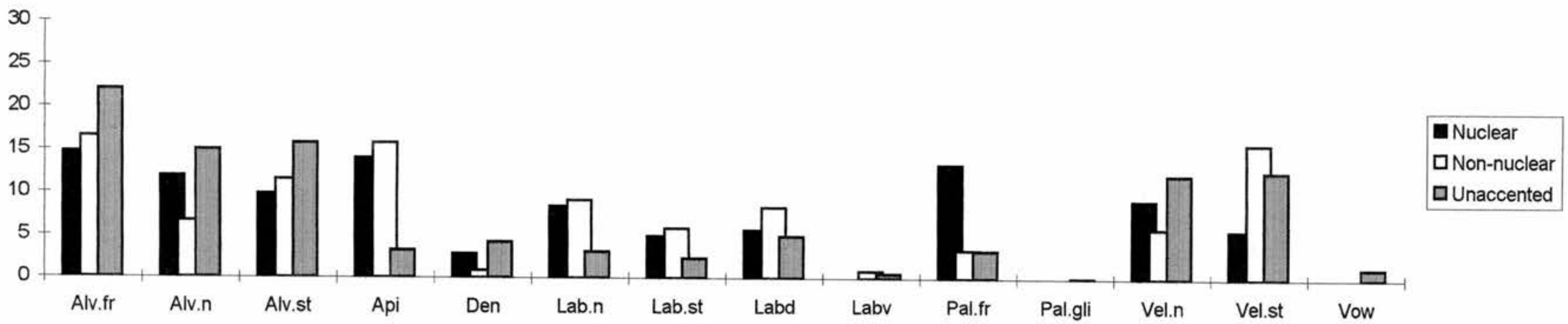
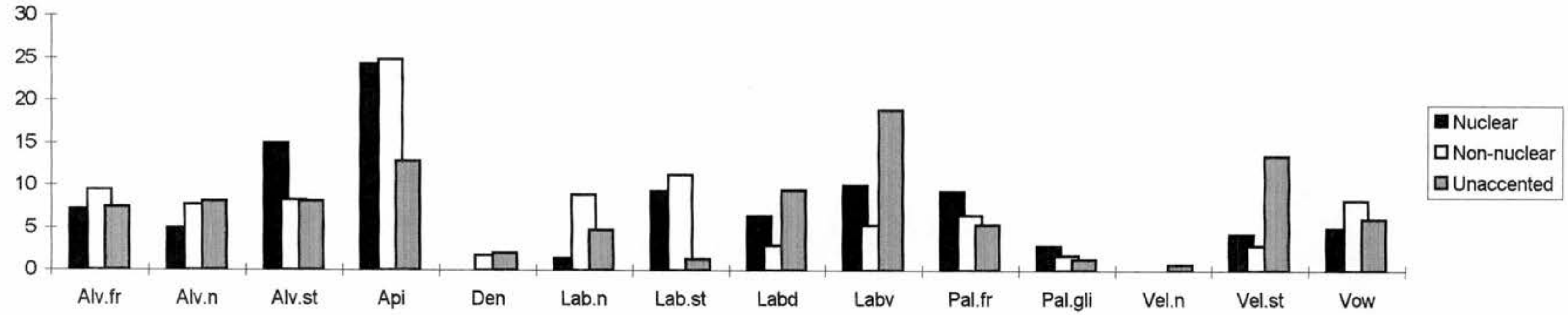


Figure 6.18: Proportional distribution of preceding and following contexts for nuclear and non-nuclear accented and unaccented /ε/ tokens

6.18a: preceding contexts



6.18b: following contexts

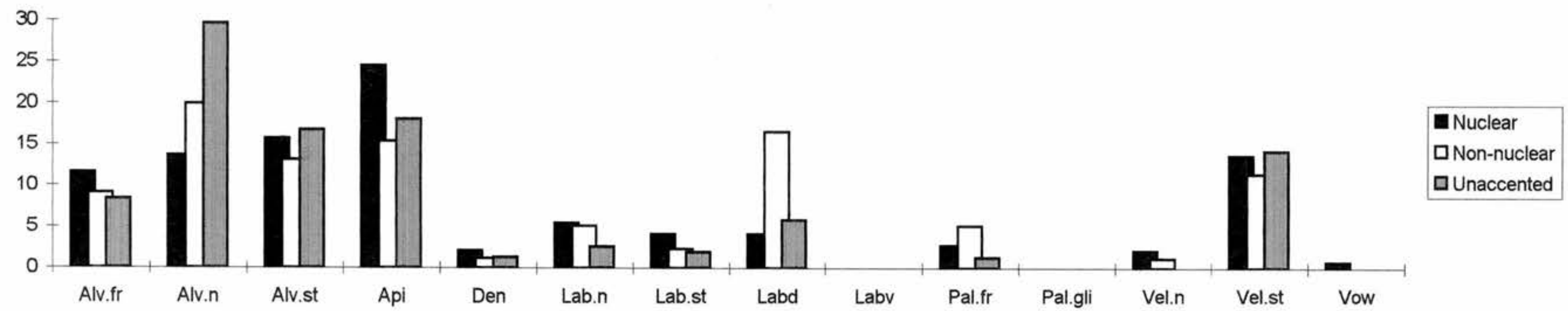
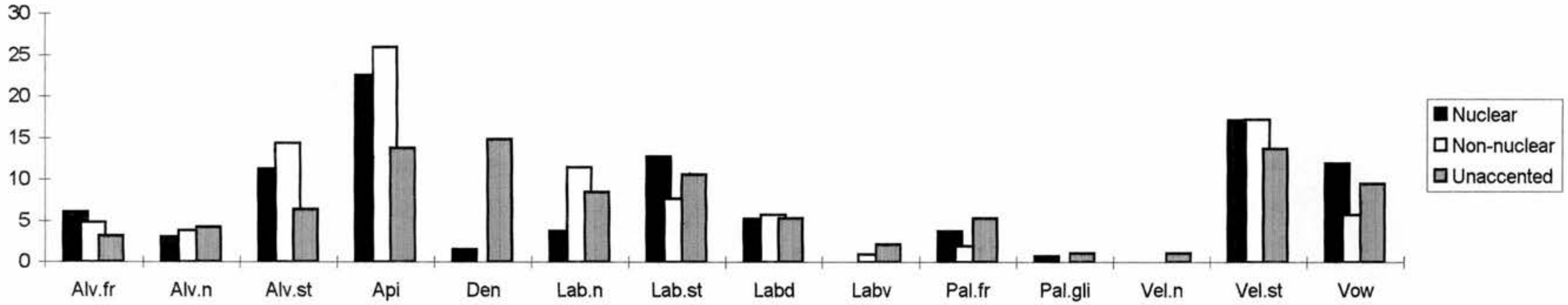


Figure 6.19: Proportional distribution of preceding and following contexts for nuclear and non-nuclear accented and unaccented /a/ tokens

6.19a: preceding contexts



6.19b: following contexts

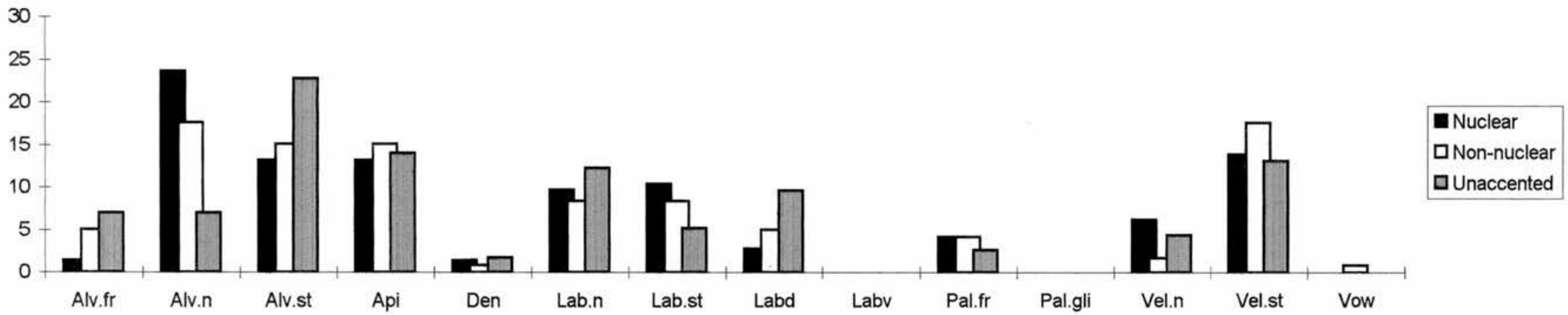
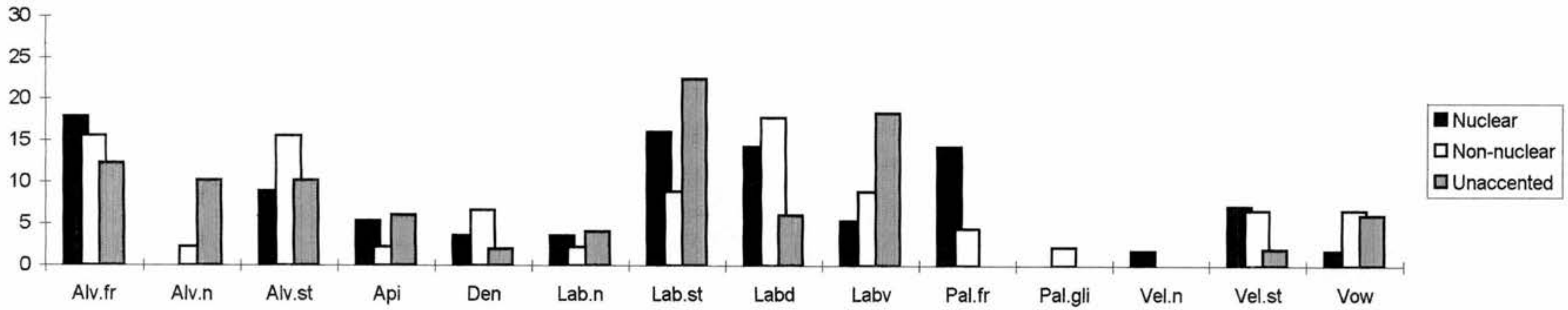


Figure 6.20: Proportional distribution of preceding and following contexts for nuclear and non-nuclear accented and unaccented /ʒ/ tokens

6.20a: preceding contexts



6.20b: following contexts

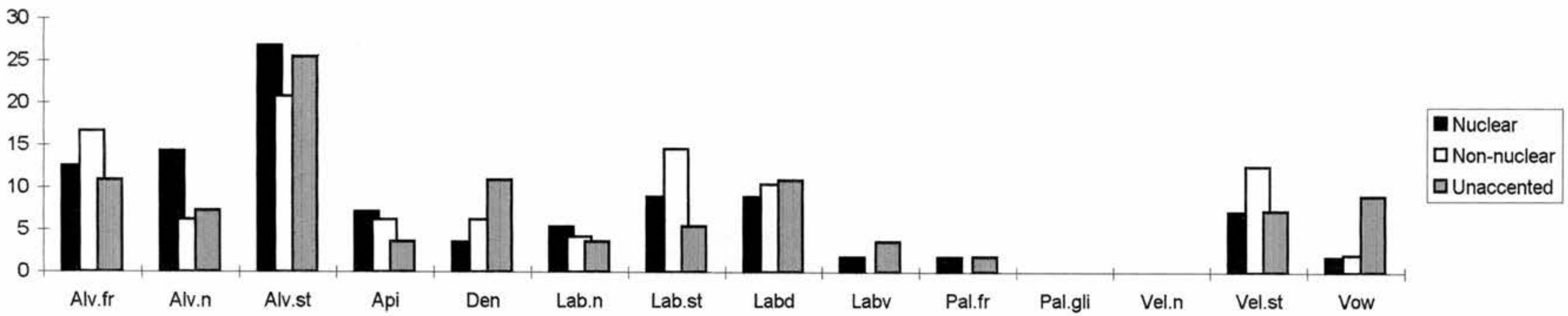
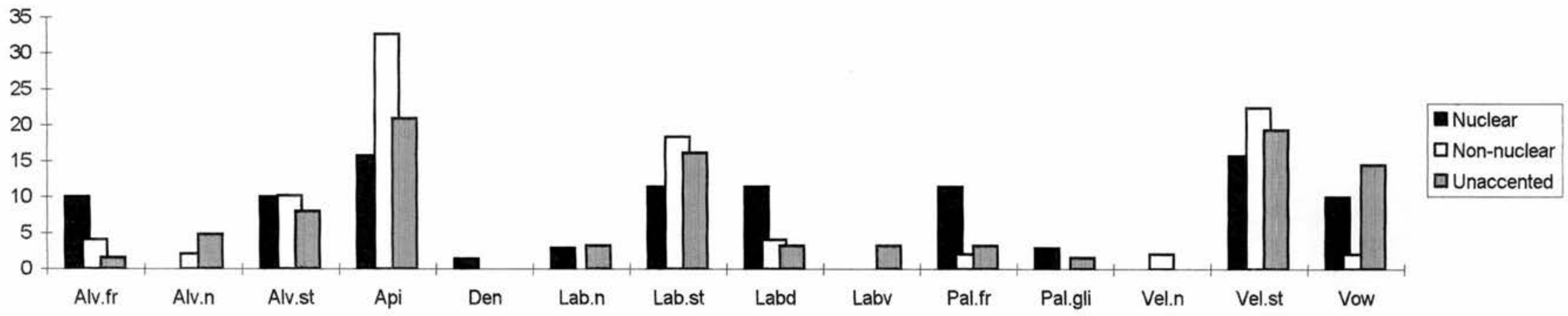


Figure 6.21: Proportional distribution of preceding and following contexts for nuclear and non-nuclear accented and unaccented /ɑ/ tokens

6.21a: preceding contexts



6.21b: following contexts

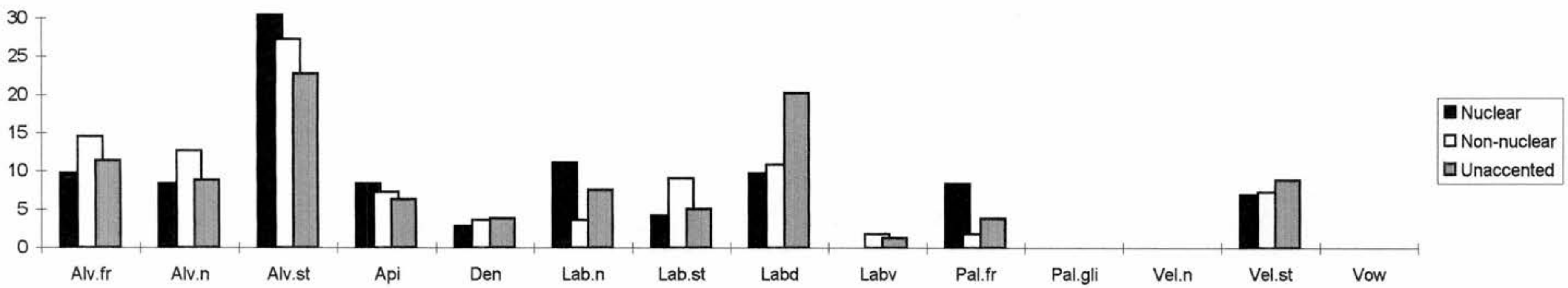
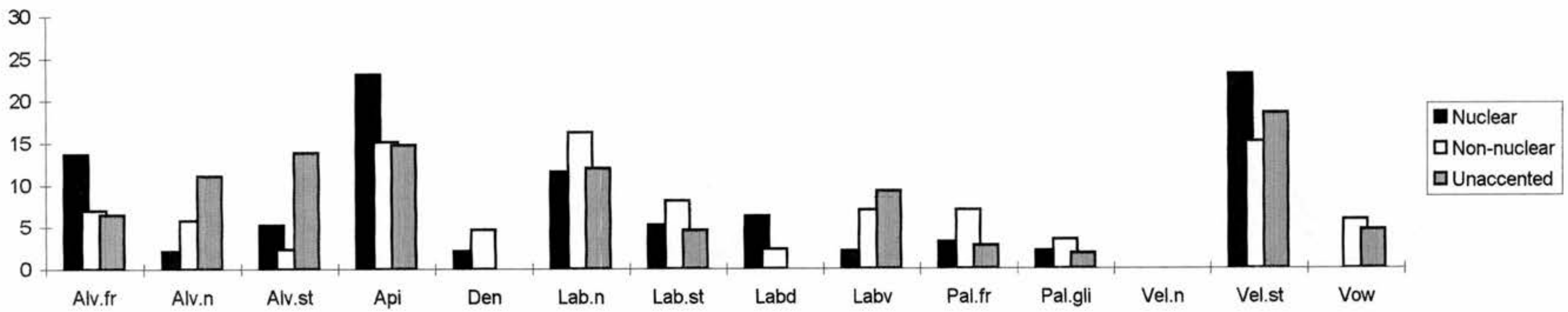


Figure 6.22: Proportional distribution of preceding and following contexts for nuclear and non-nuclear accented and unaccented /ʌ/ tokens

6.22a: preceding contexts



6.22b: following contexts

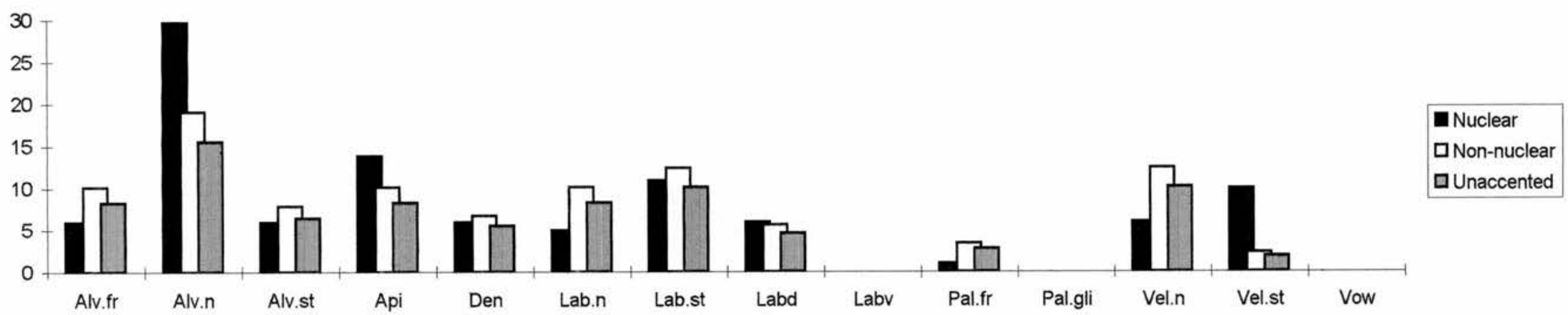
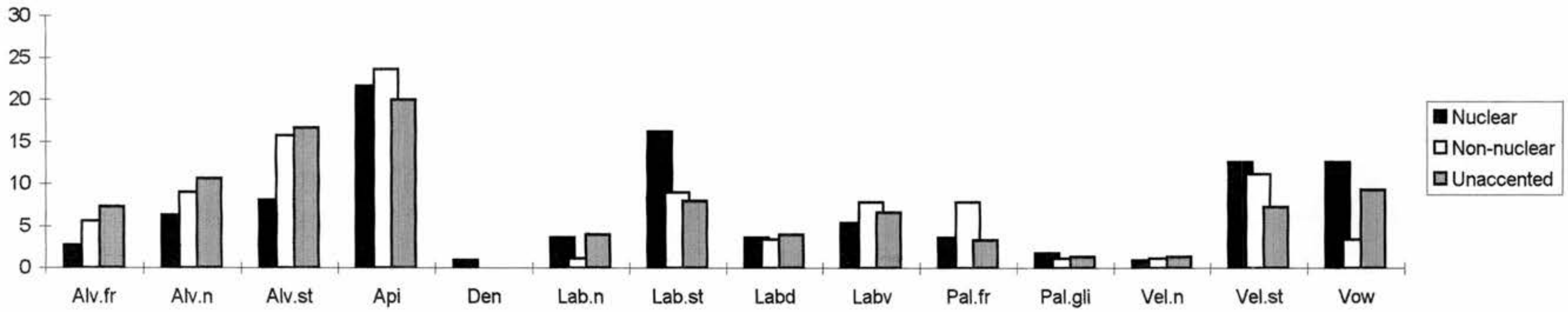


Figure 6.23: Proportional distribution of preceding and following contexts for nuclear and non-nuclear accented and unaccented /ɒ/ tokens

6.23a: preceding contexts



6.23b: following contexts

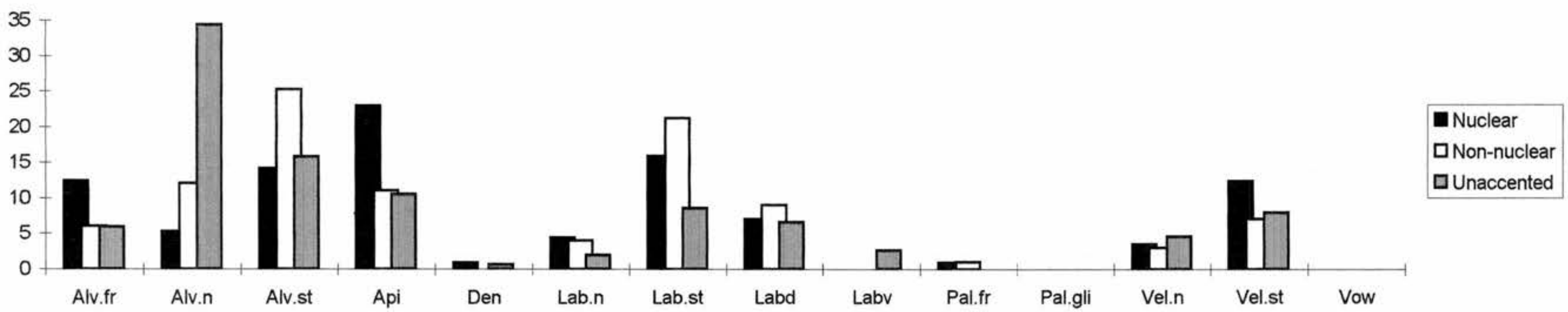
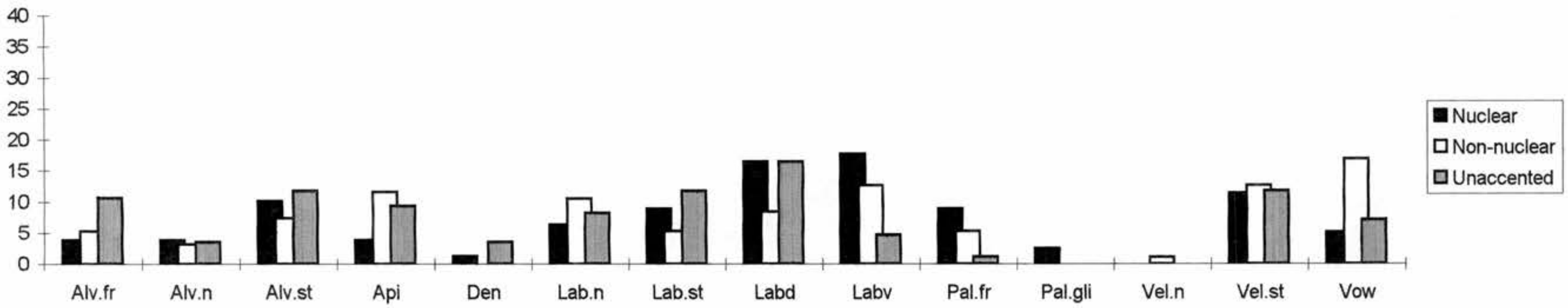


Figure 6.24: Proportional distribution of preceding and following contexts for nuclear and non-nuclear accented and unaccented /ɔ/ tokens

6.24a: preceding contexts



6.24b: following contexts

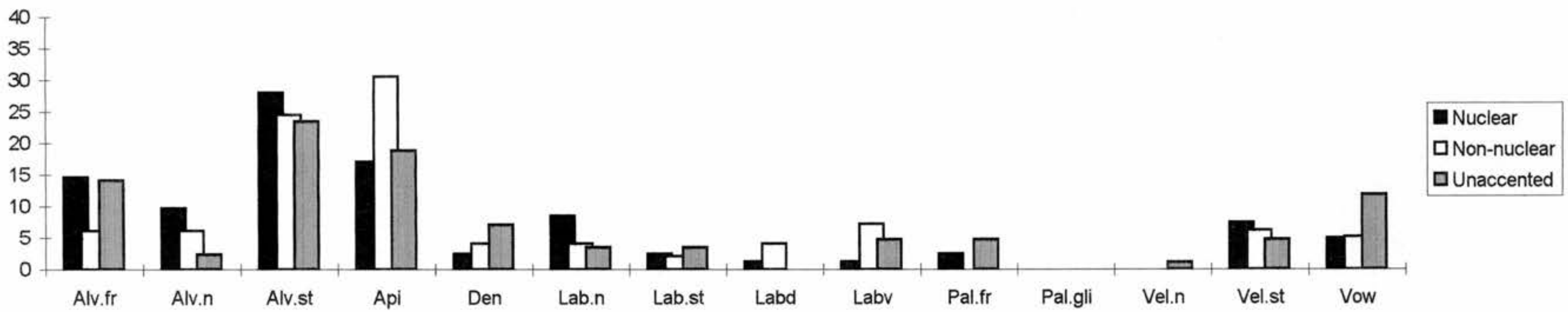
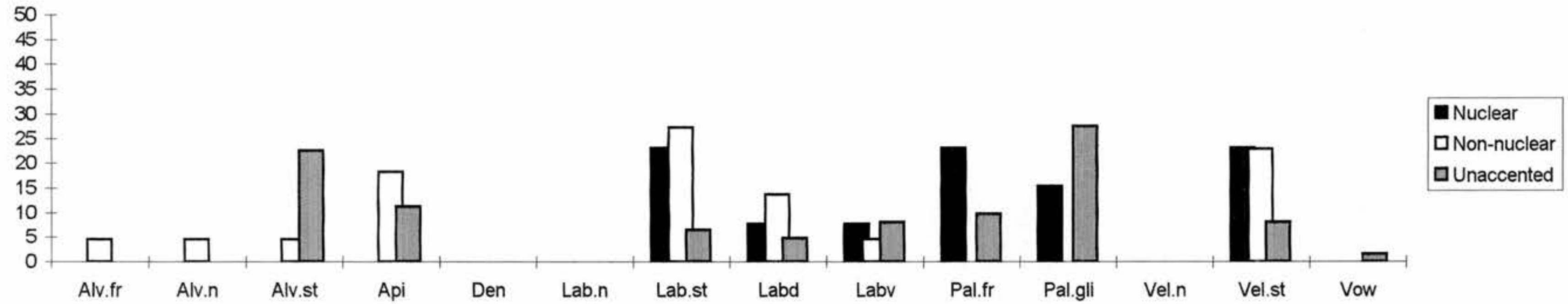


Figure 6.25: Proportional distribution of preceding and following contexts for nuclear and non-nuclear accented and unaccented /ʊ/ tokens

6.25a: preceding contexts



6.25b: following contexts

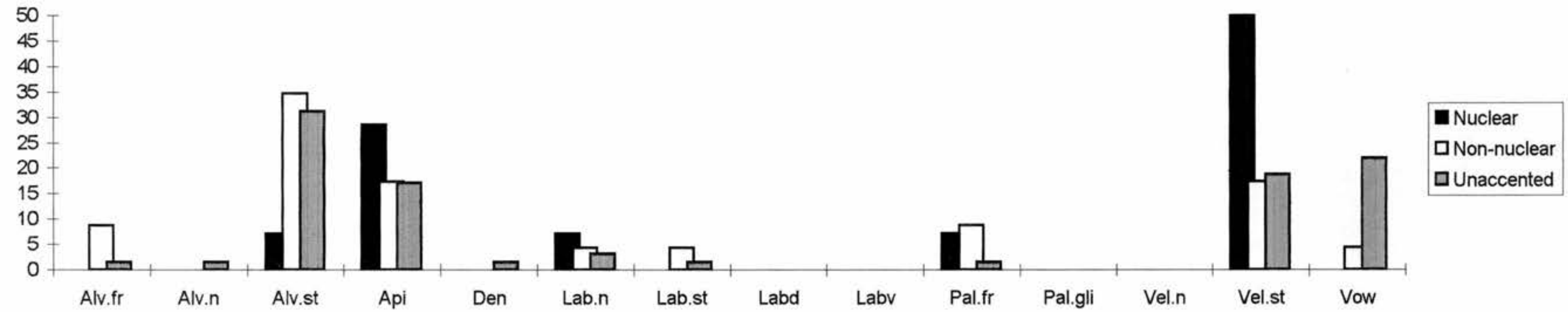
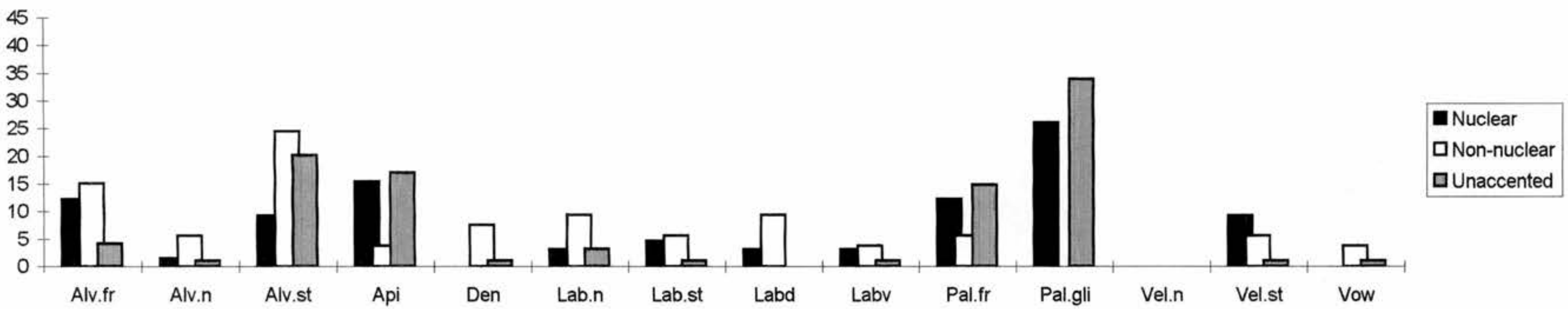


Figure 6.26: Proportional distribution of preceding and following contexts for nuclear and non-nuclear accented and unaccented /u/ tokens

6.26a: preceding contexts



6.26b: following contexts

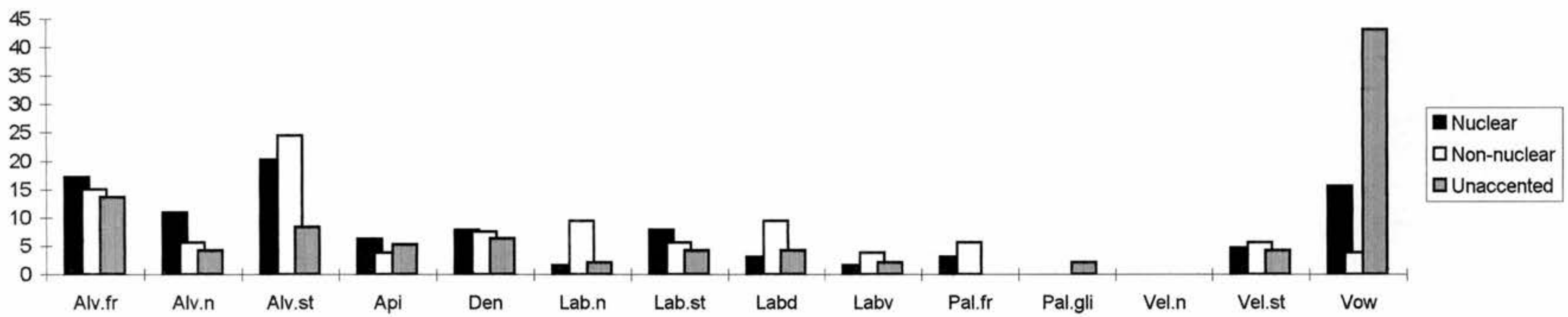


Table 6.1: Means and standard deviations for nuclear accented, non-nuclear accented and unaccented vowel tokens - F2

	Nuclear accented		*Non-nuclear accented*		*Unaccented*	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
i	2135	81	2093	76	2066	103
ɪ	1748	167	1748	178	1716	228
ɛ	1650	160	1636	142	1614	178
a	1514	86	1501	86	1492	94
ɜ	1376	51	1394	94	1372	86
ɑ	1104	59	1097	65	1120	90
ʌ	1191	72	1171	92	1220	140
ɒ	1002	85	995	85	1034	138
ɔ	809	65	838	97	885	135
ʊ	1153	185	1217	186	1402	305
u	1606	196	1629	198	1511	292

Table 6.2: Means and standard deviations for nuclear-accented, non-nuclear accented and unaccented vowel tokens - F1

	Nuclear accented		*Non-nuclear accented*		*Unaccented*	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
i	353	42	364	41	378	49
ɪ	456	50	440	49	440	65
ɛ	628	49	606	52	589	53
a	773	43	765	36	726	52
ɜ	565	30	540	29	557	47
ɑ	715	40	688	49	688	60
ʌ	686	50	666	55	640	72
ɒ	620	50	600	55	581	74
ɔ	468	36	466	29	473	42
ʊ	452	24	441	31	418	41
u	362	34	366	27	381	39

Table 6.3: Means and standard deviations for nuclear accented, non-nuclear accented and unaccented vowel tokens - duration (msec)

	Nuclear accented		*Non-nuclear accented*		*Unaccented*	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
i	130	46	101	24	84	26
ɪ	78	28	68	22	56	20
ɛ	107	39	87	21	82	27
a	134	38	115	25	103	29
ɜ	175	38	136	28	130	36
ɑ	193	47	155	38	143	38
ʌ	100	31	87	25	80	22
ɒ	114	41	98	30	91	29
ɔ	174	49	132	35	126	40
ʊ	84	26	77	23	59	18
u	140	53	102	35	84	35

Table 6.4: One-way Anova results. Main effect of stress on vowel duration, F1 and F2 midpoint values

		Duration			*F1 mid*			*F2 mid*	
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	77.47	186	.0000	10.3	659	.0000	31.24	339	.0000
ɪ	54.9	276	.0000	7.82	333	.0005	3.77	335	.0241
ɛ	26.52	367	.0000	23.25	451	.0000	2.32	412	.0997
a	22.3	307	.0000	26.93	254	.0000	1.05	326	.3508
ɜ	23.1	144	.0000	6.04	111	.0032	0.54	114	.5824
ɑ	10.17	177	.0000	5.14	174	.0068	1.12	177	.3278
ʌ	15.15	256	.0000	13.75	284	.0000	6.32	246	.0021
ɒ	18.51	272	.0000	16.64	306	.0000	4.03	332	.0186
ɔ	26.14	250	.0000	0.35	210	.7070	7.06	250	.0010
ʊ	10.48	93	.0001	5.82	93	.0042	11	61	.0001
u	33.04	157	.0000	6.16	200	.0025	6.33	204	.0022

Table 6.5: Ellipse size and degree of shrinkage as a function of stress
The areas for nuclear and non-nuclear accented and unaccented vowel ellipses are given in addition to the ellipse area for combined nuclear and non-nuclear accented tokens. The difference in ellipse area between nuclear (Nuc) and non-nuclear (Non-nuc) accented tokens and between combined accented (Acc) and unaccented (Un) tokens is also shown.

	Area				*Degree of shrinkage*			
Vowel	*Nuclear*	*Non-nuclear*	*Combined*	*Unaccented*	*Nuc/Un*	*Nuc/Non-nuc*	*Non-nuc/Un*	*Acc/Un*
i	16.17	15.11	16.33	24.61	34%	7%	39%	34%
a	17.6	14.86	16.58	23.82	26%	16%	38%	30%
ɜ	8.22	15.45	13.15	22.49	63%	47%	31%	42%
ɑ	12.07	15.97	14.55	27.38	56%	24%	42%	47%
ɔ	14.54	18.81	18.19	34.78	58%	23%	46%	48%
u	38.29	31.17	35.42	55.76	33%	19%	56%	36%
ɪ	43.9	44.86	45.73	79.24	45%	2%	43%	42%
ɛ	38.45	37.11	38.93	48.99	22%	3%	24%	21%
ʌ	20.23	28.71	24.89	58.23	65%	30%	51%	57%
ɒ	26.14	27.02	27.36	65.57	40%	3%	59%	38%
ʊ	26.81	36.02	33.74	67.89	39%	26%	47%	50%

Table 6.6: Coefficient of variance values for nuclear and non-nuclear accented and unaccented tokens within each vowel category

	F1			*F2*		
Vowel	*Nuclear*	*Non-nuclear*	*Unaccented*	*Nuclear*	*Non-nuclear*	*Unaccented*
i	12.04	11.21	13.07	3.8	3.62	4.99
a	5.5	4.71	7.2	5.69	5.75	6.29
ɜ	5.25	5.42	8.45	3.72	6.71	6.24
ɑ	5.58	7.14	8.71	5.34	5.96	8.07
ɔ	7.66	6.19	8.89	8.06	11.55	15.25
u	5.31	7.01	10.27	16.05	12.18	19.3
ɪ	10.85	11.22	14.82	9.58	10.19	13.3
ɛ	7.78	8.59	9.01	9.69	8.71	11
ʌ	7.3	8.23	11.24	6.03	7.84	11.46
ɒ	8	9.1	12.75	8.44	8.58	13.37
ʊ	5.31	7.01	9.74	16.05	15.28	21.73

Table 6.7: Degree of overlap between combined nuclear and non-nuclear accented vowel tokens

Vowel	i	a	ɜ	ɑ	ɔ	u	ɪ	ɛ	ə	ʌ	ɒ
i	*					5%	6%				
a		*		1%				5%	**1%**	1%	
ɜ			*				1%	12%	**20%**	3%	
ɑ		1%		*						65%	34%
ɔ					*				**1%**	3%	12%
u	5%					*	24%		**21%**		
ɪ	6%		1%			24%	*	7%	**42%**		
ɛ		5%	12%				7%	*	**15%**	2%	
ə		**1%**	**20%**		**1%**	**21%**	**42%**	**15%**	*	**3%**	**1%**
ʌ		1%	3%	65%				2%	**3%**	*	24%
ɒ				34%	3%				**1%**	24%	*
ʊ			1%		12%	7%	8%		**35%**		

Table 6.8: Degree of overlap between unaccented vowel tokens

Vowel	i	a	ɜ	ɑ	ɔ	u	ɪ	ɛ	ə	ʌ	ɒ
i	*					33%	29%		**2%**		
a		*	3%	2%				11%	**4%**	15%	
ɜ		3%	*			1%	5%	33%	**28%**	40%	7%
ɑ		2%		*	1%				**1%**	72%	54%
ɔ				1%	*	2%			**1%**	2%	21%
u	33%		1%		1%	*	44%		**35%**	1%	4%
ɪ	29%		5%			44%	*	15%	**64%**	2%	1%
ɛ		11%	33%				15%	*	**27%**	15%	4%
ə	**2%**	**4%**	**28%**	**1%**	**1%**	**35%**	**64%**	**27%**	*	**18%**	**9%**
ʌ		15%	40%	72%	2%	1%	2%	15%	**18%**	*	60%
ɒ			7%	54%	21%	4%	1%	4%	**9%**	60%	*
ʊ	18%		3%	1%	12%	73%	61%	1%	**58%**	3%	8%

Table 6.9: F2 Regression results for combined nuclear and non-nuclear accented and unaccented vowel tokens

	Accented						Unaccented								
	R^2	*F-value*	*df.*	β			R^2	*F-value*	*df.*	β			*Difference between groups*		
				lhs	*rhs*	*dur*				*lhs*	*rhs*	*dur*	*F-value*	*df.*	*p-value*
i	**.5021**	**66.21**	197	.37	.37	.50	**.5422**	**180.03**	456	.56	.30	.41	**10.31**	653	.0000
ɪ	**.8578**	**480.49**	239	.50	.64	.19	**.9107**	**3546.30**	1043	.59	.47	.11	**10.84**	1282	.0000
ɛ	**.6861**	**219.26**	301	.53	.55	.19	**.7203**	**122.78**	143	.51	.52	.11	**2.19**	444	.0690
a	**.3169**	**36.04**	233	.42	.40	.11	**.5261**	**32.57**	88	.63	.39	**.08**	**3.17**	321	.0142
ɜ	**.6374**	**56.25**	96	.65	.42	**.03**	**.6243**	**23.82**	43	.73	.29	**.03**	**1.64**	139	.1686
ɑ	**.0885**	**3.69**	114	.19	**-.17**	**-.17**	**.1760**	**4.13**	58	.32	**-.06**	-.28	**1.32**	172	.2636
ʌ	**.5011**	**59.26**	177	.68	.34	**-.06**	**.6732**	**70.73**	103	.75	.40	-.21	**7.50**	280	.0000
ɔ	**.4408**	**43.89**	167	.57	**.01**	-.43	**.2050**	**6.70**	78	.30	**.10**	-.44	**2.45**	245	.0471
ɒ	**.3189**	**30.43**	195	.52	**.06**	-.37	**.6230**	**79.33**	144	.60	.27	-.48	**14.60**	339	.0000
ʊ	**.8801**	**75.86**	31	.77	.33	-.18	**.9219**	**224.19**	57	.53	.58	-.13	**6.78**	88	.0001
u	**.6236**	**61.29**	111	.58	.41	**-.01**	**.8747**	**207.08**	89	.46	.64	**-.03**	**9.23**	200	.0000

Table 6.10: F2 Regression results for nuclear and non-nuclear accented vowel tokens

	Nuclear						Non-nuclear								
	R^2	*F-value*	*df.*	β			R^2	*F-value*	*df.*	β			*Difference between groups*		
				lhs	*rhs*	*dur*				*lhs*	*rhs*	*dur*	*F-value*	*df.*	*p-value*
i	**.4429**	**26.23**	99	.30	.35	.54	**.5292**	**35.22**	94	.47	.42	.29	1.50	193	.2033
ɪ	**.8052**	**175.03**	127	.49	.60	.23	**.9227**	**429.65**	108	.51	.68	.15	4.71	235	.0011
ɛ	**.6906**	**98.96**	133	.52	.59	**.00**	**.6944**	**124.24**	164	.54	.52	.23	1.72	297	.1447
a	**.3067**	**19.02**	129	.40	.38	**.13**	**.3368**	**16.93**	100	.45	.44	**.09**	0.45	229	.7754
ɜ	**.6114**	**27.28**	52	.61	.52	**.00**	**.6808**	**28.44**	40	.70	.34	**-.03**	1.65	92	.1683
ɑ	**.1193**	***2.94**	65	**.18**	**-.21**	**-.18**	***.0716**	**1.16**	45	**.18**	**-.14**	**-.17**	0.14	110	.9671
ʌ	**.4618**	**26.03**	91	.64	.33	**.18**	**.5536**	**33.90**	82	.75	.37	-.19	2.09	173	.0844
ɔ	**.3119**	**10.88**	72	.52	**-.07**	-.36	**.5308**	**34.32**	91	.61	**.10**	-.47	3.43	163	.0102
ɒ	**.2841**	**14.02**	106	.49	**.09**	-.32	**.4364**	**21.94**	85	.60	**.07**	-.51	2.74	191	.0302
ʊ	**.9318**	**40.97**	9	.85	**.20**	**-.05**	**.8917**	**49.40**	18	.80	.34	-.26	2.06	27	.1147
u	**.5670**	**25.79**	59	.50	.51	**.06**	**.7556**	**49.46**	48	.72	.29	**-.07**	2.15	107	.0798

Table 6.11: F1 Regression results for combined nuclear and non-nuclear accented and unaccented vowel tokens

	Accented						*Unaccented*						*Difference between groups*		
	R^2	*F-value*	*df.*	β			R^2	*F-value*	*df.*	β			*F-value*	*df.*	*p-value*
				lhs	*rhs*	*dur*				*lhs*	*rhs*	*dur*			
i	**.2286**	**19.36**	196	.27	.34	-.18	**.4629**	**131.56**	458	.53	.37	-.15	**3.59**	654	.0066
ɪ	**.4359**	**64.15**	249	.28	.55	**.03**	**.6744**	**752.46**	1090	.54	.44	.07	**9.70**	1339	.0000
ɛ	**.3668**	**58.51**	303	.34	.47	.36	**.4367**	**36.95**	143	.36	.44	.35	**1.59**	446	.1753
a	**.2090**	**20.34**	231	.45	.15	.38	**.5396**	**34.38**	88	.55	.29	.64	**11.26**	319	.0000
ɜ	**.4513**	**26.32**	96	.32	.40	.49	**.3411**	**7.42**	43	**.27**	**.17**	.42	**.84**	139	.5007
ɑ	**.1086**	**4.59**	113	.26	.19	**.10**	**.2279**	**5.51**	56	.41	**.09**	.37	**2.39**	169	.0531
ʌ	**.3382**	**29.98**	176	.33	.30	.48	**.5213**	**37.39**	103	.45	.37	.44	**4.00**	279	.0036
ɔ	**.0353**	**2.04**	167	.15	**.11**	**-.03**	**.1503**	**4.60**	78	.35	**.15**	**.07**	**1.57**	245	.1844
ɒ	**.1497**	**11.45**	195	.31	.21	.29	**.4719**	**42.60**	143	.43	.42	.23	**6.29**	338	.0001
u	**.1612**	**7.05**	110	**.15**	.35	**-.08**	**.6667**	**59.35**	89	.41	.55	**-.09**	**6.97**	199	.0000
ʊ	**.6449**	**18.77**	31	.59	.36	.28	**.3186**	**8.88**	57	**.15**	.51	**.01**	**4.27**	88	.0033

Table 6.12: F1 Regression results for nuclear and non-nuclear accented vowel tokens

	Nuclear							Non-nuclear									
	R^2	*F-value*	*df.*	*p-value*	β			R^2	*F-value*	*df.*	*p-value*	β			*Difference between groups*		
					lhs	*rhs*	*dur*					*lhs*	*rhs*	*dur*	*F-value*	*df.*	*p-value*
i	**.2445**	**10.57**	98	.0000	.24	.40	**-.17**	**.1970**	**7.69**	94	.0001	.31	.27	**-.08**	.76	192	.5528
ɪ	**.4196**	**32.30**	134	.0000	.23	.56	**.04**	**.4869**	**35.11**	111	.0000	.43	.53	**.09**	3.62	245	.0070
ɛ	**.3635**	**25.70**	135	.0000	.38	.47	.40	**.3322**	**27.19**	164	.0000	.29	.49	.28	.91	299	.4613
a	**.2168**	**11.72**	127	.0000	.52	**.15**	.34	**.2467**	**10.91**	100	.0000	.36	.19	.49	1.72	227	.1462
ɜ	**.2261**	**4.97**	51	.0042	.28	.28	.34	**.5430**	**16.24**	41	.0000	.43	.48	.43	1.82	92	.1320
ɑ	**.2046**	**5.66**	66	.0016	.33	**.16**	**-.15**	**.0976**	**1.55**	43	.0000	**.15**	**.26**	**.26**	3.98	109	.0048
ʌ	**.4306**	**22.94**	91	.0000	.29	.29	.57	**.2218**	**7.69**	81	.0000	.37	.34	.32	.78	172	.5393
ɔ	**.0138**	**0.34**	72	.7988	**.09**	**.08**	**-.03**	**.1002**	**3.38**	91	.0217	.28	**.12**	**-.05**	.56	163	.6924
ɒ	**.1684**	**7.16**	106	.0002	.26	.28	.36	**.1169**	**3.76**	85	.0139	.33	**.14**	**.13**	2.58	191	.0386
ʊ	**.5031**	**3.04**	9	.0855	**.26**	**.33**	.58	**.8498**	**33.95**	18	.0000	.80	.21	.25	4.08	27	.0103
u	**.1281**	**2.89**	59	.0429	**.09**	.33	**.03**	**.3372**	**8.14**	48	.0002	**.11**	.55	**-.04**	.89	107	.4744

Table 6.13: Difference between groups - F2

	Nuclear accent vs unaccented			*Non-nuclear accent vs unaccented*		
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	8.61	555	.0000	4.43	550	.0016
ɪ	13.47	1170	.0000	1.03	1151	.3919
ɛ	.78	276	.5418	3.45	307	.0089
a	3.52	217	.0083	1.4	188	.2362
ɜ	3.19	95	.0167	.40	84	.8098
ɑ	1.05	123	.3823	.53	103	.7123
ʌ	6.38	194	.0001	2.84	185	.0255
ɒ	13.22	250	.0000	4.79	144	.0010
ɔ	1.88	150	.1177	3.18	169	.0151
ʊ	3.87	66	.0070	4.82	75	.0016
u	3.9	148	.0048	11.16	137	.0000

Table 6.14: Difference between groups - F1

	Nuclear accent vs unaccented			*Non-nuclear accent vs unaccented*		
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	2.59	556	.0357	2.66	552	.0321
ɪ	13.35	1224	.0000	0.2	1201	.9377
ɛ	2.11	278	.0799	0.86	307	.4904
a	7.38	215	.0000	7.94	188	.0000
ɜ	.89	94	.4721	0.75	84	.5607
ɑ	4.58	122	.0018	0.97	99	.4268
ʌ	3.88	194	.0047	1.87	184	.1170
ɒ	5.75	249	.0002	3.7	228	.0062
ɔ	1.43	150	.2259	0.67	169	.6116
ʊ	2.01	66	.1039	3.36	75	.0138
u	4.81	148	.0011	2.28	137	.0642

Table 6.15: Two-way Anova results for F2. Main effect of stress and C^1/C^2 place of articulation

	Stress			C^1			*Interaction*			*Stress*			C^2			*Interaction*		
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	29.97	1, 79	.0000	11.59	7, 80	.0000	5.12	7, 81	.0001	4.25	1, 15	.0571	6.93	7, 473	.0000	2.64	7, 473	.0109
ɪ	4.81	1, 173	.0297	4.36	7, 1256	.0001	6.62	7, 177	.0000	46.42	1, 7	.0003	11.77	7, 4	.0156	2.43	7, 4	.2042
ɛ	6.77	1, 406	.0096	20.03	7, 26	.0000	4.85	7, 28	.0011	3.91	1, 437	.0485	53.19	6, 437	.0000	2.28	6, 437	.0350
a	1.84	1, 282	.1756	9.25	7, 282	.0000	2.56	7, 282	.0143	2.25	1, 310	.1344	19.07	5, 71	.0000	0.57	5, 310	.7240
ɜ	0.61	1, 122	.4378	4.31	5, 122	.0012	1.32	5, 15	.3085	0.61	1, 22	.4378	4.31	5, 122	.0012	1.32	5, 15	.3085
ɑ	1.35	1, 5	.4527	0.31	1, 5	.8684	0.16	1, 5	.9448	.10	1, 157	.7469	0.8	5, 157	.5524	0.49	5, 157	.7855
ʌ	4.94	1, 253	.0271	22.92	5, 253	.0000	1.76	5, 253	.1216	4.9	1, 274	.0276	1.6	6, 274	.1460	1.14	6, 274	.3385
ɒ	4.20	1, 45	.0462	5.68	6, 47	.0002	.99	6, 46	.4443	.01	1, 333	.9313	1.21	6, 38	.3221	6.09	6, 40	.0001
ɔ	9.71	1, 211	.0011	8.77	7, 211	.0000	2.04	7, 211	.0517	2.04	1, 215	.1552	1.24	6, 215	.2860	1.38	6, 215	.2249
ʊ	29.01	1, 71	.0000	1.84	4, 71	.1308	3.36	4, 71	.0142	29.01	1, 71	.0000	1.84	4, 71	.1308	3.36	4, 71	.0142
u	12.82	1, 190	.0004	34.49	5, 16	.0000	4.61	5, 22	.0050	5.61	1, 141	.0192	11.93	7, 119	.0000	0.8	7, 141	.5912

Table 6.16: Two-way Anova results for F1. Main effect of stress and C^1/C^2 manner of articulation

	Stress			C^1			*Interaction*			*Stress*			C^2			*Interaction*		
	F-value	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*	*F-value*	*df.*	*p-value*
i	9.24	1, 636	.0024	13.14	4, 121	.0000	3.24	4, 121	.0144	2.69	1, 650	.1009	7.1	5, 74	.0000	3.69	5, 75	.0048
ɪ	9.28	1, 113	.0029	39.27	4, 114	.0000	2.61	4, 1257	.0340	2.49	1, 25	.1271	40.55	4, 10	.0000	4.96	4, 1323	.0006
ɛ	28.41	1, 414	.0000	3.34	4, 414	.0105	1.81	4, 414	.1266	28	1, 445	.0000	6.72	3, 445	.0002	2.64	3, 445	.0488
a	13.15	1, 286	.0000	1.7	4, 286	.1491	0.3	4, 286	.8764	29.8	1, 56	.0000	4.2	3, 318	.0062	0.84	3, 318	.4741
ɜ	0	1, 30	.7957	1.27	4, 130	.2847	0.21	4, 22	.9281	0	1, 127	.9719	5.42	2, 127	.0055	2.55	2, 127	.0824
ɑ	.06	1, 151	.8034	0.94	4, 151	.4408	0.07	4, 151	.9914	.02	1, 160	.8929	12.94	2, 160	.0000	3.78	2, 160	.0249
ʌ	16.87	1, 268	.0001	5.44	4, 268	.0003	0.14	4, 192	.9660	9	1, 279	.0029	6.19	3, 143	.0006	1.44	3, 279	.2323
ɒ	15.67	1, 109	.0001	1.09	4, 305	.3634	3.56	4, 305	.0075	23.63	1, 80	.0000	2.50	3, 80	.0652	2.49	3, 80	.0661
ɔ	1.32	1, 21	.2643	3.89	4, 217	.0045	1.24	4, 217	.2936	0.84	1, 10	.3813	3.19	4, 11	.0573	1.29	4, 224	.2742
u	2.5	1, 15	.1347	3.35	4, 197	.0112	2	4, 14	.1506	1.79	1, 19	.1968	10.48	4, 147	.0000	1.63	4, 147	.1689
ʊ	4.55	1, 86	.0358	3.84	3, 86	.1463	0.82	3, 86	.4867	10.17	1, 74	.0021	1.78	3, 74	.1586	2.12	3, 74	.1045

Table 6.17: Distribution of preceding contexts for nuclear (Nuc) and non-nuclear (Acc) accented and unaccented (Un) vowel tokens

Context	i			ɪ			ɛ			a			ɜ			ɑ			ʌ			ɒ			ɔ			ʊ			u		
	Nuc	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*
s z	10	7	34	19	17	152	10	16	11	8	5	3	10	7	6	7	2	1	13	6	7	3	5	11	3	5	9	0	1	0	8	8	4
n	3	7	45	3	1	70	7	13	12	4	4	4	0	1	5	0	1	3	2	5	12	7	8	16	3	3	3	0	1	0	1	3	1
t d	15	14	61	16	15	262	21	14	12	15	15	6	5	7	5	7	5	5	5	2	15	9	14	25	8	7	10	0	1	14	6	13	19
l r	40	31	186	38	19	171	34	42	19	30	27	13	3	1	3	11	16	13	22	13	16	24	21	30	3	11	8	0	4	7	10	2	16
ð θ	2	1	62	7	13	36	0	3	3	2	0	14	2	3	1	1	0	0	2	4	0	1	0	0	1	0	3	0	0	0	0	4	1
m	3	4	16	7	3	41	2	15	7	5	12	8	2	1	2	2	0	2	11	14	13	4	1	6	5	10	7	0	0	0	2	5	3
p b	16	7	40	13	14	54	13	19	2	17	8	10	9	4	11	8	9	10	5	7	5	18	8	12	7	5	10	3	6	4	3	3	1
f v	5	5	19	8	13	51	9	5	14	7	6	5	8	8	3	8	2	2	6	2	0	4	3	6	13	8	14	1	3	3	2	5	0
w	5	5	21	11	7	85	14	9	28	0	1	2	3	4	9	0	0	2	2	6	10	6	7	10	14	12	4	1	1	5	2	2	1
ʃ ʒ	3	8	20	6	10	49	13	11	8	5	2	5	8	2	0	8	1	2	3	6	3	4	7	5	7	5	1	3	0	6	8	3	14
j	0	0	0	0	0	1	4	3	2	1	0	1	0	1	0	2	0	1	2	3	2	2	1	2	2	0	0	2	0	17	17	0	32
ŋ	0	0	1	0	4	22	0	0	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1	2	0	1	0	0	0	0	0	0	0
k g	4	2	9	8	0	38	6	5	20	23	18	13	4	3	1	11	11	12	22	13	20	14	10	11	9	12	10	3	5	5	6	3	1
Vowel	4	8	4	3	1	76	7	14	9	16	6	9	1	3	3	7	1	9	0	5	5	14	3	14	4	16	6	0	0	1	0	2	1
ʔ h	2	2	20	3	1	16	6	4	5	9	6	18	0	2	8	0	0	0	6	2	1	3	6	2	4	3	1	1	1	3	1	2	2
Silence	0	6	1	1	3	25	1	3	3	2	9	5	0	1	0	0	0	0	0	1	0	0	4	0	2	0	2	0	0	0	0	0	0

Table 6.18: Distribution of following contexts for nuclear (Nuc) and non-nuclear accented (Acc) and unaccented (Un) vowel tokens

Context	i			ɪ			ɛ			a			ɜ		
	Nuc	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*
s z	23	13	59	21	20	252	17	16	13	2	6	8	7	8	6
n	13	6	21	17	8	171	20	35	46	34	21	8	8	3	4
t d	19	38	53	14	14	180	23	23	26	19	18	26	15	10	14
l r	6	8	23	20	19	37	36	27	28	19	18	16	4	3	2
ð θ	3	1	10	4	1	48	3	2	2	2	1	2	2	3	6
m	3	5	16	12	11	35	8	9	4	14	10	14	3	2	2
p b	7	9	41	7	7	26	6	4	3	15	10	6	5	7	3
f v	12	14	20	8	10	56	6	29	9	4	6	11	5	5	6
w	2	1	21	0	1	6	0	0	0	0	0	0	1	0	2
ʃ ʒ	4	1	4	19	4	37	4	9	2	6	5	3	1	0	1
j	0	0	3	0	0	1	0	0	0	0	0	0	0	0	0
ŋ	0	1	4	13	7	138	3	2	0	9	2	5	0	0	0
k g	9	7	39	8	19	143	20	20	22	20	21	15	4	6	4
Vowel	5	2	169	0	0	14	1	0	0	0	1	0	1	1	5
ʔ h	0	1	46	0	0	4	0	0	1	0	0	3	0	0	0
Silence	6	0	10	0	0	2	0	0	0	0	0	0	0	0	2

Context	ɑ			ʌ			ɒ			ɔ			ʊ			u		
	Nuc	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*	*Nuc*	*Acc*	*Un*
s z	7	8	9	6	9	9	14	6	9	12	6	12	0	2	1	11	8	13
n	6	7	7	30	17	18	6	12	52	8	6	2	0	0	1	7	3	4
t d	22	15	18	6	7	18	16	25	24	23	24	20	1	8	20	13	13	8
l r	6	4	5	14	9	8	26	11	16	14	30	16	4	4	11	4	2	5
ð θ	2	2	3	6	6	8	1	0	1	2	4	6	0	0	1	5	4	6
m	8	2	6	5	9	18	5	4	3	7	4	3	1	1	2	1	5	2
p b	3	5	4	11	11	10	18	21	13	2	2	3	0	1	1	5	3	4
f v	7	6	16	6	5	13	8	9	10	1	4	0	0	0	0	2	5	4
w	0	1	1	0	0	0	0	0	4	1	7	4	0	0	0	1	2	2
ʃ ʒ	6	1	3	1	3	1	1	1	0	2	0	4	1	2	1	2	3	0
j	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
ŋ	0	0	0	6	11	3	4	3	7	0	0	1	0	0	0	0	0	0
k g	5	4	7	10	2	3	14	7	12	6	6	4	7	4	12	3	3	4
Vowel	0	0	0	0	0	0	0	0	0	4	5	10	0	1	14	10	2	41
ʔ h	1	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	2	1
Silence	0	0	0	0	0	0	1	0	0	3	0	0	0	0	1	2	0	0

Chapter 7
Discussion and Conclusions

This chapter reviews the main findings of this thesis and considers the theoretical implications these have for an account of vowel reduction in English. In particular, the results are examined for evidence to support an account of vowel reduction in terms of phonetic underspecification. The results are discussed first with respect to the three main aims of this thesis. These were, firstly, to obtain a measure of schwa's variability in comparison with the full vowels and to determine what evidence there is to support the hypothesis that schwa is unspecified for tongue and jaw position; secondly, to determine whether the full vowels vary significantly with respect to the degree of context sensitivity they show and thirdly, to establish whether sententially unstressed vowels are more susceptible to coarticulatory effects than their sententially stressed counterparts.

In the second part of the discussion, the systematicity and directionality of the observed coarticulatory effects are discussed more generally with respect to the theoretical implications they have for different models of coarticulation and for the invariance issue in speech production. A review of the advantages and disadvantages of a corpus-based study as opposed to an experimentally controlled paradigm is also given here. The chapter concludes with a summary of the main points of interest arising from this thesis.

7.1 Phonetic underspecification of schwa

As predicted, in the present data, schwa displays the greatest overall variability and the highest overall level of context-dependency of all the vowels. The near-maximal context-dependency for schwa along F2 strongly supports the hypothesis that schwa is completely unspecified for tongue position. The data also indicates that the second formant values observed for schwa are independently derived by principles of phonetic implementation as opposed to being supplied by phonological "fill-in" rules. As Keating (1988) predicts, unspecified segments are likely to show a dynamic, transitional quality along the dimension for which they lack specification. This reflects their "interpolation through" by the trajectory between adjacent specified segments (see section 2.3.2.2). The linearity of schwa second formant trajectories, the wide range in schwa F2 midpoint value, the comparable directionality of effects observed for schwa and the interaction of context effects through schwa all indicate its phonetic transparency and interpolation through.

These results concur with findings reported by Choi (1992) for short vowels in Marshallese. In Marshallese, the short vowels are fully specified for tongue height but are unspecified for front/back tongue position. In an acoustic investigation, Choi also demonstrates that the observed F2 values for these vowels are determined by the values associated with the characteristics of the adjacent consonants and by the interpolation function governing the consonant-to-consonant trajectory:

> F2 at the vowel midpoint was found to vary primarily as a function of consonantal secondary articulation type. While there was a weak effect on F2 associated with vowel category, this was attributed to variation in constriction size (i.e. the phonemic vowel height contrast) and not to an inherent front/back vocalic specification. The data also showed robust consonant-to-consonant coarticulation, providing further evidence that the vowels lack a F2 target. These results ... all suggest that the vowels of Marshallese can be modeled without reference to an F2 target in the underspecification paradigm. (p. 108)

In the present data, the results for schwa are the more striking given that they indicate underspecification along both principle dimensions of vowel quality: front/backness and height. While it is conceded that some degree of jaw opening is inherent to schwa's [-consonantal, +syllabic] specification, beyond being characterised by an open approximation as opposed to a stricture of complete closure, there is no evidence in the present data of a unique F1 target for schwa. The relatively high F1 context-dependency displayed by schwa in comparison with the full vowels (excepting /ɪ/), the linear F1 trajectories for schwa and an interaction of F1 context effects through schwa indicate that it is also highly unspecified for jaw position.

The results for schwa thus provide empirical support for Keating's (1988) Underspecification Hypothesis. Evidence that schwa is targetless and can occupy almost any position in the vowel space depending on context also argues against the traditional concept of vowel reduction as an independent process of centralisation. It is also consistent with the view that vowel reduction represents a means of economising on articulatory effort (Van Bergem, 1995) insofar as the straight-line interpolation between adjacent context segments reflects minimal articulatory effort.

7.2 Differences between the full vowels with regard to inherent variability

The present data also provide evidence in support of Keating's (1990) proposal that segments may show varying degrees of specification along a given dimension. According to Keating's window model of coarticulation, segments are characterised by the full range of contextual variability they exhibit. Segments with a precise or narrow specification for a given feature show less overall variability along the corresponding phonetic dimension(s) than segments which are less narrowly specified. The present results show significant differences between vowels with respect to overall variability and degree of context-dependency along both F1 and F2.

Comparable levels of context-dependency along F2 observed for the mid-high, lax vowels /ɪ/ and /ʊ/ (for both accented and unaccented tokens) to that displayed by schwa suggests that these vowels are also unspecified for front/back tongue position. Both vowels display a comparable range in F2 midpoint value to schwa and similarly little difference in range of F2 midpoint compared with range in F2 onset and offset value, indicating interpolation through. An interaction of context effects through the vowel is also evidence of their phonetic transparency along this dimension.

The differences in mean F2 value between the three vowels (/ə/ 1447 Hz, /ɪ/ 1716 Hz, /ʊ/ 1325 Hz) may be attributed to differences in the frequency distribution of contexts. For example, for /ɪ/ there is a relatively higher proportion of palatal and velar contexts and a lower proportion of labial/labio-dental contexts than for schwa (see Figure A 1). The higher proportion of velar contexts, in particular, has the effect of raising the grand mean for /ɪ/. For example, /ɪ/ tokens is the context of a preceding or following velar show a mean F2 value of 1884 Hz. The mean value for /ɪ/ tokens in the context of a preceding or following labial (where the opposing context is any consonant other than a palatal or velar) show a mean F2 value of 1580 Hz (see Table A 1.b).

Overall, /ɪ/ also shows a comparable level of F1 variability and degree of F1 context-dependency to schwa which suggests that, like schwa, it is also highly unspecified for tongue height. However, the stressed/unstressed comparison of the data shows a considerable reduction in degree of F1 context-dependency, thus indicating a narrower tongue-height specification, for accented /ɪ/ tokens relative to unaccented /ɪ/ tokens. Generally low F1 context-dependency for /ʊ/ also indicates an inherent specification for tongue height and/or lip rounding. Complete underspecification, as in the case of schwa, would not be expected for these vowels given that, unlike schwa, they may both occur in stressed position.

The remaining vowels show varying degrees of context-dependency. The ranking of vowels from least to most context-dependent along F1 shows greater variation as a function of sentence stress than the ranking of vowels in terms of F2 context-dependency. However, as in the case of F2, schwa and /ɪ/ show the highest degree of context-dependency for F1 within each accent condition while the mid-low, back vowels show the lowest F1 context-dependency. With the exception of /i/, all vowels predictably show a lesser degree of context-dependency along F1 than along F2 (see also section 7.5.1).

Following the mid-high, lax vowels /ɪ/ and /ʊ/, the lax, front vowel /ɛ/, the long, central vowel /ɜ/, the tense vowel /ɯ/ and the lax, back vowel /ʌ/, in descending order, show the highest levels of F2 context dependency in the accented condition. In the unaccented condition, /ɯ/ shows a higher degree of context-dependency than /ɛ, ɜ/ and /ʌ/. The back vowels /ɑ/, /ɒ/ and /ɔ/ show the least context-dependency along F2 in the accented condition, followed by the tense, front vowel /i/ and the long, lax vowel /a/. In the unaccented condition, /i/ shows a lesser degree of context-dependency than /ɒ/.

It is proposed here that this pattern of results reflects a continuum of underspecification which, broadly speaking, ranges from the tense (or more peripheral) vowels which may be thought of as the most narrowly specified and hence least contextually variable vowels, to the less narrowly specified and more contextually variable lax (or more central) vowels, to schwa, which, being completely unspecified, shows maximal context-dependency.

7.2.1 Inherent vowel duration

To some extent, the observed hierarchy of robustness parallels the hierarchy of inherent vowel duration. The inherently long vowels /ɑ, ɔ, a/ are among the most robust vowels. The lax vowel /ɒ/ shows greater stability than the inherently shorter lax vowels /ʌ/ and /ɛ/. The mid-high, lax vowels /ɪ/ and /ʊ/ which are characterised by the shortest intrinsic durations after schwa, also show the greatest context-dependency after schwa. However, there are also discrepancies between the two rankings of vowels. For example, in the accented condition /ɯ/ shows higher context-dependency than /ɒ/ and /ʌ/ despite showing significantly longer durations. Similarly, in the unaccented condition, /ɯ/ shows a greater degree of context-dependency than /ɛ/, /ʌ/ and /i/ although it shows no significant durational difference. In both accent conditions, /ɛ/ displays a higher R^2 value than /ʌ/ despite being of a similar duration. Similarly, the long, central vowel /ɜ/, which shows comparable durations to /ɑ/ and /ɔ/ also shows a higher degree of context-dependency. Conversely, /i/ which shows the same duration as /ɛ/ and /ʌ/ in the unaccented condition also shows a considerably lower level of F2 context-dependency.

As in the case of F2, the long vowels /ɑ/, /ɔ/ and /a/ show relatively low context-dependency along F1 in both the accented and unaccented condition while /ɪ/ shows the highest F1 context-dependency. However, /i/ shows greater context-dependency than /a/, /ɛ/ and /ʌ/ in the accented condition despite showing similar durations to /a/ and significantly longer durations than /ɛ/ or /ʌ/. Similarly, accented /ɜ/ and /u/ tokens show greater context-dependency than accented /i/ tokens although /u/ is characterised by comparable durations to /i/ and /ɜ/ by significantly longer durations than /i/.

The pattern of results across stress conditions also presents conflicting evidence regarding the potential correlation between stability and duration. The tense vowel /u/ which shows the largest relative decrease in duration for unaccented relative to accented tokens also shows the largest relative increase in amount of context-dependency. Similarly, in the case of /ɒ/, the relatively large decrease in duration with a decrease in stress is also accompanied by a relatively large increase in level of context-dependency. However, the tense vowel /i/ which shows the largest relative decrease in duration across accent conditions following /u/, shows a relatively small increase in amount of context-dependency. /ɪ/ and /ʊ/ also display a relatively small increase in context-dependency with reduced stress despite showing a relatively large decrease in duration. The long vowels /ɑ, ɜ, ɔ/ show no significant increase in degree of context-dependency with a decrease in stress although they all show a significant decrease in duration.

The discrepancies between the two rankings of vowels and the apparent lack of correlation between relative decrease in duration on the one hand and relative increase in context-dependency on the other with reduced stress, indicate that the present pattern of results reflects factors other than inherent vowel duration. An alternative possibility which would appear to give a more complete account of the data, is offered by Stevens' (1972, 1989) quantal theory of speech production.

7.2.2 Inherent acoustic properties of vowels

According to Stevens, there are regions in the vocal tract which show quantal relations between articulation and acoustics. Articulatory perturbations in these regions have a minimal effect on the acoustic signal whereas at the boundaries of these regions, articulatory variations of the same magnitude can produce significant acoustic variability. These regions also correspond to the areas which produce the most distinctive acoustic patterns:

> ... there are some articulatory states or configurations or gestures that give rise to well-defined patterns of auditory response in the listener, such that these patterns are not strongly sensitive to small perturbations or inaccuracies in the articulation. These patterns are distinctive in the sense that if some articulatory parameter crosses over a threshold region there will be a significant change or a qualitative shift in the auditory response. The multi-dimensional space that depicts acoustic-articulatory regions or auditory-articulatory relations, rather than showing continuous and monotonic variation, exhibits quantal attributes characterised by rapid changes in state over some region and less abrupt variations or greater stability over other regions. (Stevens, 1989, p. 5)

These regions are identified as the places of articulation which produce similar resonances in the front and back oral cavities and which thus give rise to well defined spectral peaks (Fant, 1956). Accordingly, greater acoustic stability is predicted for the point vowels /i, ɑ, u/, described as the limiting articulations of the vowel triangle (Lieberman & Blumstein, 1988), than for other vowels. These vowels, represent the extremes in vowel quality insofar as the oral and pharyngeal tubes are maximally expanded and/or maximally constricted in their production (Fant, 1960; Stevens, 1972).

The high, front vowel /i/ is characterised by a wide pharyngeal region and a narrow constriction in the palatal region. This configuration results in a high frequency spectral peak due to the convergence of F2 and F3. The low, back vowel /ɑ/ is formed by a narrow constriction in the lower pharyngeal region, resulting in a relatively large cross-sectional area in the anterior part of the oral tract. In this case, F1 and F2 converge to produce a spectral peak at about 1000 Hz. The high, back vowel /u/ is also characterised by a large oral cavity although in this case, the major constriction is in the upper pharyngeal region. However, the additional lengthening of the vocal tract through lip-rounding serves to lower F2 with the result that F2 and F1 converge to produce a low frequency spectral peak.

The non-point vowels which are articulated at the boundaries of the quantal regions and which are also characterised by more widely separated formants are expected to show greater acoustic variability:

> When the cross-sectional area of the constriction is increased or decreased, keeping the constriction position fixed at one of the stable regions, the formants tend to change monotonically. That is, there is not a well-defined range of constriction sizes for which the formant frequencies achieve maximum or minimum values and thus are relatively insensitive to constriction size. (p. 15)

The present data accords with Stevens' predictions in showing greater acoustic stability for the more peripheral compared with the more central vowels. The low, back vowel /ɑ/ displays relatively low overall variability and the least context-dependency along F2. Along F1, it shows the least context dependency after /ɔ/. The high, front vowel /i/ also displays relatively low F2 context-dependency. The relatively high F1 variability observed for /i/ is also consistent with Stevens' predictions: "we note ... that F1 varies monotonically with constriction position and with constriction size for the .. (/i/) .. configuration" (p.12).

The high overall variability observed for /u/ in the present data may be attributed to its fronted realisation. With a more anterior place of articulation, the frequency of F2 increases, increasing the distance between F1 and F2 and therefore also increasing the sensitivity of F2 to anterior-posterior perturbations in tongue-body position. High F2 variability for /u/ may also be attributed to its specification for lip-rounding. According to Stevens (1989), variation in the extent of the lip-rounding gesture produces gradual and monotonic changes in formant frequencies. Variability is further increased through a high tolerance for articulatory imprecision. Given that there are physiological constraints on degree of lip-rounding, it is possible "to reproduce a particular degree of rounding without requiring a great deal of precision in the degree of excitation of the appropriate muscles" (p. 16).

The high degree of stability observed for the other rounded vowels /ɔ/ and, to a lesser extent, /ɒ/, reflects the relatively close proximity of F1 and F2 achieved through the lowering effect on F2 of the lip rounding gesture in combination with a posterior tongue-body constriction:

> ...for vocal tract shapes that are rounded ... there is a range of constriction positions that will yield a value of F2 that is low, is relatively stable, and is close to F1. These positions are achieved by manipulating the tongue body to a backed position to form a constriction in the region of the soft palate or upper pharynx. For these configurations, F1 is lower than the values that could be reached if there were no constriction at the lips. (p. 14)

In the present data, the low, front vowel /a/ is also relatively robust compared with the mid, front vowels and the mid, back vowel /ʌ/. The low jaw position and narrowing in the posterior part of the vocal tract during production of this vowel also results in a relatively close proximity of F1 and F2 compared with the other front vowels. The greater stability of /a/ relative to /ʌ/ may reflect the fact that, in the case of /a/, F1 is close to its maximum. While F1 and F2 are generally more proximal in the case of /ʌ/, neither F1 nor F2 is close to its respective maximum or minimum.

To date, there has been little empirical evidence reported in support of quantal theory predictions. Pisoni (1982) addresses the question of inherent vowel variability with respect to predictions generated by the quantal theory. However, he reports no clear pattern of variances across the eight vowels in his study. Pisoni's findings may be due in part to the fact that he evaluates vowel variability in terms of the standard deviations of the formant frequencies without controlling for the correlation between standard deviations and formant frequency value. His results may also reflect methodological problems and the artificial nature of the speech material used. This comprised isolated productions of the vowels /i, ɪ, ɛ, æ, ʌ, ɑ, ɔ, u/ by two speakers in response to synthetic vowels. Pisoni notes that the vowel tokens were, in general, longer than the synthetic tokens and showed no differentiation between vowel categories in terms of inherent vowel duration.

However, in accord with the present results and thus also with quantal predictions, Stevens & House (1963) report a greater shift in F2 value as a function of consonantal context for lax vowels than for tense vowels. They also report that, in their data, the long vowels /i, æ, ɑ/ show the least variability in F2 value while the rounded vowels /u/ and /ʊ/ show the highest variability. The results of psychoacoustic tests also consistently show higher recognition rates for the point vowels and for tense vowels generally, than for lax vowels (Peterson & Barney, 1952; Strange et al.,1976; Fowler & Shankweiler, 1978; Strange, 1989b), thereby also indicating greater acoustic stability for these vowels.

To summarise the findings for the present data; in addition to the relatively high acoustic stability observed for the long vowels /a, ɑ, ɔ/ and the relatively low stability observed for the short vowels /ɪ, ɛ, ʌ, ʊ/, quantal theory also accounts for the high acoustic stability observed for /i/ despite its relatively short duration in the non-nuclear accented and unaccented condition. It also accounts for the relatively high acoustic stability displayed by /ɒ/ within each accent condition compared with the other lax or short vowels. It also offers an explanation for the relatively high F1 variability displayed by /i/ and for the high F2 context-dependency observed for /u/ in each accent condition which, likewise, cannot be explained in terms of inherent duration.

The observed hierarchy of robustness thus arguably reflects the inherent acoustic properties of vowels rather than their inherent duration. However, while the corner vowels /i, a, ɑ/ and the back, rounded vowels /ɒ/ and /ɔ/ may be more stable acoustically than the more central vowels, the question remains as to whether they are generally more robust, that is, whether they are also more resistant to coarticulatory effects at the articulatory level. It is possible that greater articulatory variability is tolerated for these vowels precisely on account of their quantal properties and the non-monotonic relation between articulation and acoustics. Thus, greater acoustic stability does not necessarily imply greater articulatory stability and generally less context-sensitivity.

There is some evidence in the literature that articulatory programming "takes into account the non-linear or "quantal" relationships between articulation and acoustics" (Perkell & Nelson, 1985, p. 1895). Perkell & Nelson (1985) examine the relative positioning of points on the tongue surface using x-ray microbeam data for the vowels /i/ and /a/. For these vowels, variations in the degree of vocal-tract constriction produce larger percentage changes in the area function and vowel formant frequencies than variations in the location of the maximal constriction. In view of this, Perkell & Nelson hypothesise that greater articulatory variability will be tolerated with respect to the location of maximal constriction than with respect to the degree of maximal constriction. They examined multiple repetitions of /i/ and /a/ tokens produced by three speakers of American English in a variety of context and stress conditions. In support of their hypothesis, the results showed greater precision "in the positioning of dorsal tongue points near the place of maximal constriction ... in a direction perpendicular to the vocal-tract midline than in a direction parallel to the midline" (p. 1895).

Given that Perkell & Nelson (1985) only examine data for the quantal vowels /i/ and /a/, the question remains as to how quantal vowels compare with non-quantal vowels in terms of constraints on articulatory variability. In order to do justice to this question, it would be necessary to perform a simultaneous examination of articulatory and acoustic vowel data. Such an examination is beyond the scope of this thesis. However, notwithstanding this limitation, the present acoustic data is consistent with predictions made in the literature on the basis of articulatory, perceptual and language-specific phonological constraints. These are discussed below.

7.2.3 Articulatory constraints on variability

In Chapter 2 it was suggested that the more peripheral vowels may be expected to show a lesser overall degree of context-dependency than the more central vowels on the grounds that they show a lesser overall degree of gestural compatibility with context. Recasens (1991) defines gestural compatibility in terms of the degree of constraint on articulatory activity. He proposes that during vowel production, only those parts of the tongue body directly involved in the vocalic constriction are under active control, leaving those parts not directly involved free to coarticulate with upcoming segments. Thus, he predicts greater anticipatory coarticulation for back vowels than for front vowels on the assumption that during their production only the back of the tongue is subject to constraint (see section 2.3.2.1 for a full discussion).

The present vowel data shows a marked distinction between full and reduced vowels with respect to relative degree of anticipatory coarticulation. Schwa and /ɪ/ both display a high degree of differentiation in F2 midpoint value as a function of following consonant place of articulation in contrast to the other vowels which display relatively little systematic variation in F2 midpoint value as a function of following consonantal place characteristics. However, there is no distinction between front and back vowels in this respect (see section 5.4.2.1). The long, central vowel /ɜ/ also displays relatively little variation as a function of following consonant place of articulation. Assuming that, contrary to Recasens' proposal, the entire tongue body is under active control during the production of full vowels irrespective of their front/back specification, degree of constraint might then be defined in terms of the location of the maximal constriction and the degree of constriction. The front vowel /i/ and the back vowels /ɑ, ɒ, ɔ/, in representing the extremes in front/back tongue position, and the low vowels /a/ and /ɑ/ in representing maximal jaw opening, are thus arguably subject to a greater degree of articulatory constraint than the more central vowels.

Clearly, at the individual token level, different vowel-consonant and consonant-vowel sequences are likely to show varying degrees of coarticulation depending on a variety of factors including degree of stress, duration and the specific combination of gestures. In the case of antagonistic vowel and consonant gestures, there is also the question of whether the articulatory requirements for the consonant take precedence over those for the vowel or vice versa (see Recasens, 1991). However, given that there are a greater number of contexts which involve the tongue trajectory passing through the central area of the vowel space than through the more peripheral regions, it would seem reasonable to assume that the lesser overall degree of context-dependency observed for the peripheral vowels reflects a lesser overall degree of gestural compatibility.

7.2.4 Phonological constraints on variability

The apparent discrepancy between Recasens' (1991) observations for vowels in Catalan and the present results may, to some extent, reflect language-specific phonological constraints on variability. Manuel (1990) proposes that there are "output constraints" on the amount of contextual variability a given phone may exhibit. These are determined, in part, by the phonemic inventory of the language such that tighter output constraints apply in situations where a high degree of coarticulation would lead to a major loss of distinctiveness. Thus, she argues that languages with a large number of vowels might be expected to show less anticipatory coarticulation than languages which have smaller vowel inventories. In support of this hypothesis, Manuel cites findings from an earlier study by Manuel & Krakow (1984) which reports greater anticipatory vowel-to-vowel coarticulation for English than for Swahili and Shona, two languages which have five vowel systems. In addition to the number of vowels within a given system, Manuel (1990) claims that the way in which the vowels are distributed is also likely to determine the restrictions on contextual variability. Thus, greater anticipatory coarticulation might be expected in those areas of the vowel space which are relatively uncrowded compared with other areas.

In section 5.4.2.1, it was suggested that the relatively low degree of anticipatory coarticulation observed for back vowels in the present study may be partly attributable to contextual limitations since there are relatively few examples of following palatal contexts. However, given that the back vowel area is more crowded in English than in Catalan, the discrepancy in results may also reflect tighter output constraints for the English vowels.

7.2.5 Perceptual constraints on variability: "anchor" vowels

Greater acoustic stability for the vowels /i, a, ɑ/ and /ɔ/ also accords with the view that the corner vowels serve to delimit and normalise the vowel space for a given speaker (Joos, 1948; Nearey, 1978; Henton, 1983; Liberman, 1984). Thus, greater acoustic stability observed for these vowels may also reflect tighter perceptual constraints on production.

Overlap in vowel formant frequency value occurs as a function of both within-speaker variability due chiefly to phonetic context and between-speaker variability due chiefly to differences in vocal tract size. To date, vowel recognition models have largely focused on the problem of between speaker differences. Essentially two types of vowel normalisation procedure have been proposed to deal with this. These may be referred to as "intrinsic" and "extrinsic" methods (Ainsworth, 1975).

Intrinsic methods exploit the inter-segmental relational properties of vowels, in particular, the relations between the fundamental and formant frequencies (Syrdal, 1984, Syrdal & Gopal, 1986; Miller, 1984, 1989). According to this approach, all the information required for vowel identification is contained within the vowel itself. In contrast, extrinsic methods use information that is distributed across a speaker's entire vowel system such as relative vowel duration and relative formant frequency value (Joos, 1948, Ladefoged & Broadbent, 1957; Gerstman, 1968; Nearey, 1978).

Two assumptions are implicit to the intrinsic approach. As outlined by Nearey (1978) and cited in Nearey (1989), the first of these are that there are universal constraints on the shape of the vowel space for any given speaker, i.e. the "vowel triangle". The second is that there are also constraints on the nature of speaker differences such as would permit a uniform scaling of formant frequencies. Given these constraints, "certain vowels could not overlap in the F1 x F2 space ... [and] ... might then serve to "calibrate" the rest of the system" (Nearey, 1989, p. 2092).

For the present SBS speaker, this function is arguably fulfilled by the corner vowels /i, a, ɑ/ and the back, rounded vowel /ɔ/. (Owing to the fronted nature of /u/, either /o/ or /ɔ/ is most likely to fulfill the role of non-low, back anchor vowel.) The distribution of all vowel tokens in the present data pooled across all contexts and stress conditions is shown in Figures 7.1:3. The set of long vowels, the set of short vowels and schwa are plotted separately.

As Figure 7.1 demonstrates, the distribution of /i/, /a/, /ɑ/ and /ɔ/ tokens delimit a clear "vowel triangle". With the exception of some overlap between /i/ and /u/ due to the fronting of /u/, the long vowels also show a distinct clustering. In contrast, as Figure 7.2 demonstrates, there is considerable overlap between the short vowels /ɪ, ɛ, ʌ, ɒ, ʊ/. In Figure 7.3, the distribution of schwa tokens is plotted. While these cover a larger area than the tokens for any other vowel, the majority of tokens are largely confined to the region occupied by the short vowels. There is comparatively little overlap between schwa and the corner or 'anchor' vowels (see also section 6.3.2).

7.2.6 Summary

The hierarchy of robustness observed for vowels in the present data (where robustness is defined in terms of overall degree of context-dependency), shows some correspondence with the hierarchy of inherent vowel duration. The inherently long vowels /ɑ, ɔ, a/ are among the most robust while the inherently short vowels /ɪ, ɛ, ʌ, ʊ/ are among the least robust. There are, however, also discrepancies between the two rankings of vowels which suggest that vowel stability cannot be accounted for solely in terms of inherent duration. In accordance with Stevens' (1972, 1989) quantal theory of speech production, a more satisfactory account of the data is that the observed hierarchy of robustness reflects the inherent acoustic properties of vowels. This explains the greater acoustic stability of the peripheral vowels compared with the more central vowels and also accounts for the apparent inconsistencies between degree of context-dependency and relative vowel duration (see section 7.2.2).

Evidence that vowel quality may vary as a function of duration but independently of context also argues against an undershoot account of the data (see section 5.5). The fact that the durational effect is largest in the case of /i/ and /ɑ, ɒ, ɔ/ which represent the extreme upper and lower limits in F2 value, is consistent with a gestural overlap account in which vowels are characterised in terms of regions rather than as points in articulatory and acoustic space. In representing the extreme front/back tongue positions, these vowels arguably have more scope for variation (i.e. centralisation) without being subject to overlap.

Although it is not testable in this thesis, greater acoustic stability for the peripheral vowels generally compared with the more central vowels arguably reflects a lesser overall degree of compatibility with context and subsequently a lesser overall degree of gestural overlap. This assumes that gestural compatibility is defined in terms of the location and degree of maximal constriction. It also presupposes that there are constraints on the articulators to attain the required positions. One rationale for tighter constraints on variability is that these vowels serve as the anchor or reference vowels by which the entire vowel space for a given speaker is calibrated (see section 7.2.5). In view of these considerations, it is proposed here that the peripheral vowels /i, a, ɑ, ɒ, ɔ/ are more narrowly specified for tongue and jaw position than the more central vowels /ɪ, ɛ, ʌ, ʊ, u/ where this implies greater articulatory and/or acoustic stability.

7.3 Role of stress

The results of the stressed/unstressed comparison of full vowel tokens concur with the results reported for schwa insofar as they also provide evidence to support an account of vowel reduction in terms of contextual assimilation. In general, unaccented tokens show greater variability and greater overall context-dependency than their accented counterparts. There is also a tendency for degree of context-dependency to decrease linearly with an increase in degree of sentence stress. Greater context sensitivity and, in the case of the peripheral vowels, generally less extreme formant values for unaccented relative to accented vowel tokens are also consistent with the view that vowel reduction represents a means of economising on articulatory effort.

Lindblom (1983, 1990, 1992) suggests that the speaker varies his pronunciation along a continuum of hyper- and hypo-speech which ranges from maximally explicit articulations (hyper-speech) to extreme reductions and under articulation (hypo-speech). The degree of explicitness with which a given word or phrase is produced is determined by the listener's requirements for intelligibility on the one hand and the speaker's desire for economy of articulatory effort on the other. Lindblom proposes that this balance of forces varies according to short-term fluctuations in the amount of non-signal information that is available to the listener as well as in response to situational variables. He suggests that the speaker is tacitly aware of the amount of signal independent information the listener has access to and will adapt his speech accordingly. As the amount of explicit physical information required by the listener decreases "... the motor system tends to default to a low-cost form of behaviour" (Lindblom, 1990, p. 413).

The use of sentence stress is one way in which information may be highlighted (Bolinger, 1975). Van Bergem (1990) also proposes that stressed vowels serve as "anchor" points in the word recognition process. Lexically stressed vowels play a greater role in determining "the uniqueness of words in terms of their phonemic composition" (Van Bergem, 1995, p. 23) as demonstrated in studies by Carter (1987) and Altman & Carter (1989). This is due to the greater variety of vowels which occur in stressed syllables. The high perceptual salience of stressed vowels has also been demonstrated experimentally (Bond, 1981) as has the low perceptual salience of vowels in unstressed syllables (Van Bergem, 1995). Thus, accepting Lindbloms's (1983, 1990) contention that the speaker strives to maintain a balance between intelligibility and economy of articulatory effort, reduced stress may be considered a means of economising on articulatory effort in less informative parts of an utterance.

The fact that the same general pattern of variability across vowel categories is evident for both accented and unaccented tokens supports the view that the observed hierarchy of robustness reflects inherent differences between vowels. The lack of a significant increase in context-dependency along F2 in the case of /ɑ/ and /ɔ/ and the relatively small increase in F2 context-dependency in the case of /i/ is consistent with the view that these vowels are subject to tighter constraints on variability than the more central vowels. Conversely, the comparatively small stress effect for F2 observed for /ɪ/ and /ʊ/ and hence high context-dependency for accented as well as unaccented tokens, further testifies to the weak specification of these vowels.

On the basis of these results, it is proposed that vowel reduction operates along two continuua in parallel, on the one hand, along a peripheral-non-peripheral continuum of inherent vowel quality and, on the other, along an stressed-unstressed continuum. This proposal is justified if one accepts that the acoustic and physiological correlates of stress consist in intensifying phonetic factors already present (Lehiste, 1970). Thus, it is proposed here that the more central vowels and unaccented vowel tokens are articulated with a lesser degree of vocal effort than the more peripheral vowels and accented vowel tokens. This results in less displacement from neutral and/or weaker constraints on variability leading to a greater degree of contextual assimilation. It is, however, important to note that the two reduction 'processes' function independently insofar as the peripheral-non-peripheral continuum reflects inherent differences in vowel quality whereas the stressed-unstressed continuum may be considered to form part of a larger, separate continuum of hyper-hypoarticulation which, broadly speaking, is style-conditioned.

7. 4 Parallels between phonetic and phonological vowel reduction

The pattern of variability observed in the present acoustic data also reflects the pattern of alternation evident in phonological vowel reduction (see section 2.7). The acoustic stability of /ɑ/ parallels the fact that this vowel only rarely alternates with schwa morpho-phonemically. The greater acoustic stability of the tense (long) vowels (excepting /u/) compared with the lax (short) vowels parallels the tense-lax alternation evident in phonological vowel reduction. The extensive morpho-phonological alternation of /ɪ/ and /ɛ/ with schwa is also paralleled by the high degree of contextual variability these vowels exhibit. The comparable level of context-dependency observed for /ɪ/ as for schwa, in particular, accords with its status as the other reduced vowel in English (Fudge, 1984).

Clearly, a more rigorous investigation of the parallels between phonetic and phonological reduction is warranted. The present observations are largely impressionistic and would benefit from a quantitative analysis of the extent of the morpho-phonemic alternation between full vowels and schwa and/or [ɪ] and between tense and lax vowels. While this is beyond the scope of this thesis, the present data are consistent with the idea that phonetic and phonological vowel reduction form part of the same historical continuum (Van Bergem, 1990). Phonetic reduction and obligatory phonological reduction may be considered to occupy the two extreme endpoints and non-obligatory phonological reduction, an intermediate position. One of the advantages of an account of vowel reduction in terms of phonetic underspecification is that it captures this relationship.

Phonetic vowel reduction is a gradual process, it is a continuous, coarticulatory effect. All vowels produced in context are subject to phonetic reduction although, as the present data indicate, the lax or more central vowels are inherently more susceptible to coarticulatory effects than the tense or more peripheral vowels. Stressed vowels also tend to show greater resistance to context effects than unstressed vowels. Non-obligatory phonological vowel reduction is traditionally classified as a discrete alternation between tense and lax vowels or between full vowels and schwa in lexically unstressed syllables (Chomsky & Halle, 1968; Liberman & Prince, 1977; Hayes, 1984). However, the fact that it is optional in its application (i.e. there are two or more pronunciation variants) and subject to style-conditioning (see section 2.7) indicates that it has not lost its phonetic motivation and therefore still qualifies as a 'gradual' process at the production level.

It is possible for a process to be both lexically discrete and yet phonetically gradual on account of the lack of isomorphy between articulation, acoustics and perception. Phonetic gradualness is not matched by perceptual gradualness. Evidence in the literature suggests that once listeners have assigned speech sounds to the phonemes in their language, they disregard acoustic differences within a phoneme category (Eimas & Miller, 1978). This is manifest in an apparently greater ability to discriminate tokens at or near a phoneme boundary than tokens well within a phoneme category (Liberman et al., 1957).

Thus, it is suggested here that non-obligatory phonological reduction reflects varying degrees of contextual assimilation (schwa being equated with complete assimilation). The speaker's ambition to produce a full vowel (and hence override coarticulatory effects) is conditioned by the intrinsic properties of words such as word stress, word class and word frequency and by stylistic considerations. Vowels in stressed syllables are pronounced more carefully than vowels in unstressed syllables because they carry a greater functional load (Bolinger, 1975). Similarly, the speaker is likely to expend less effort on function words and on words with a high frequency of occurrence than on content words or words with a low frequency of occurrence which are potentially more problematic for the listener. Minimal ambition to produce a full vowel manifests itself as a minimal expenditure of effort resulting in minimal displacement from neutral and complete contextual assimilation.

Obligatory phonological vowel reduction requires a separate treatment from non-obligatory phonological reduction and phonetic vowel reduction insofar as it is both a lexically and phonetically discrete process. Reduction is a stable part of the phonology and insensitive to changes in rate and/or care of articulation. The speaker 'aims' to produce a schwa rather than a full vowel. However, it is proposed here that the schwa which occurs in the citation form of words (the end product of obligatory phonological reduction) has the same phonetic characteristics as schwas which represent reduced full vowels (the end product of non-obligatory reduction), that is, it is targetless and completely dependent on context.

7.5 Coarticulatory effects

7.5.1 Systematicity of effects

One motivation in conducting an observational study of vowel data, was to establish how far patterns of variability and coarticulatory effects observed for highly controlled laboratory speech material would be evident in connected speech data that is largely uncontrolled with respect to factors such as segmental context and stress. Not unexpectedly, vowels in the present data display considerably higher variances in both first and second formant frequency value than is generally reported in the literature (Peterson & Barney, 1952; Stevens & House, 1963; Stevens, House & Paul, 1966; Syrdal & Gopal, 1986; Sussman, 1991 inter alia.). However, despite the lack of restriction on segmental sequences and the interaction of context effects, highly systematic coarticulatory effects are observed for vowels along both formant dimensions.

In general, the observed pattern of consonantal place effects is consistent with effects reported in the literature and the predictions made by the acoustic resonator theory (Stevens & House, 1956), such that, as consonantal place of articulation moves forward in the vocal tract, F2 locus values become increasingly lower (see section 5.3.1). The systematicity of consonant place (and manner) effects in CV and VC sequences is particularly noteworthy given that the consonant and hence the vowel onset and offset values, are also likely to reflect the influence of the next adjacent context segments.

The systematicity of coarticulatory effects in the present data has important theoretical implications with respect to an issue which has for long been a central concern in speech production research, namely, the invariance issue. The invariance issue concerns the question of how perceptual constancy is achieved in spite of considerable physical variability in the acoustic signal. Early attempts to address this question focussed on identifying invariant and static acoustic attributes at the level of the individual segment. For example, Delattre et al., (1952) attempt to characterise stop place of articulation in terms of fixed locus frequencies. Other researchers propose that invariant acoustic cues for stop place of articulation reside in the overall gross shape of the spectrum at the onset of the release burst (Stevens & Blumstein, 1979, 1981).

In contrast to the idea of absolute articulatory and acoustic invariance, Jakobsen and Fant propose that the universality and invariance of distinctive features is relational (Jakobsen & Fant, 1963). The idea that invariance may be defined relationally has recently been the focus of work by Sussman and colleagues (Sussman, 1990b; 1994; Sussman et al., 1991). Sussman (1990b) presents evidence which indicates that stop place of articulation may be classified on the basis of the relationship that exists between the value of F2 at the vowel onset/offset and the value of F2 at the vowel nucleus as measured by the "locus-equation". The locus-equation, initially formulated by Lindblom (1963), is derived by plotting the formant values at the consonant-vowel boundary against values sampled at the vowel nucleus. A regression line is then fitted to the data-points according to the following equation: $F2_i = kF2_t + c$ where $F2_i$ denotes the onset value (or initial locus) of the consonant, $F2_t$ is the vowel nucleus and k and c are constants. The slope of the regression line is given by k and represents the amount of coarticulation between locus and nucleus.

In a comparison of /bVt/, /dVt/ and /gVt/ syllables, Sussman reports a steeper slope and hence more coarticulation, for bilabials (.91) than for velars (.79) or alveolars (.54). Although he found that a single regression line could be fitted to the data points in the velar context, better fits were obtained by using separate regression lines for front and back vowel contexts. In a linear discriminant analysis using F2 onset and nucleus values as predictors, Sussman reports high categorical classification success for both bilabials and alveolars (81%) and a lower success rate for velars (68%) except when restricted to front vowel contexts (93%). However, in a further analysis using the slopes and y-intercepts obtained across sixty speakers as predictor variables, 100% percent correct categorical classification accuracy was achieved for all three places of articulation.

On the basis of these results, Sussman claims that the locus-equation provides a higher-order metric for classifying stop-place of articulation in that "it is derived over and characterises an entire stop place category" (p. 50-3). Thus, "despite extreme context-dependency of the coarticulated stop + vowel gesture, a relational form of invariance is captured when a whole phonemic category is analysed" (p. 50-5).

Sussman further claims that the existence of systematic relationships in the acoustic signal provides evidence which argues against gestural invariance theories such as the motor theory (Liberman & Mattingly, 1985) and the direct realist account of speech perception (Fowler, 1986). Liberman & Mattingly posit a special decoding mechanism which mediates phonetic perception by recovering the articulatory gesture from the acoustic signal. Fowler also maintains that information in the signal reflects the articulatory gestures. In contrast, Sussman et al., (1991) argue in favour of "a direct acoustic/auditory based utilisation of the speech waveform to recover the phonetic segment" (p. 1321). They maintain that the locus-equation provides direct evidence that phonetic categories can be acoustically represented: "if such phonological constructs are already sufficiently contrastive in their physical instantiation in the speech waveform, then a further transform to their underlying gestural form is not seen as necessary" (p. 1321).

Sussman (1994) extends the analysis of stop place of articulation to include different manner classes. He reports a similar differentiation in place of articulation for nasal stops and for fricatives as observed in the case of the oral stops. Only the approximants /w, j, l, r/ showed a unique clustering of onset values which were also relatively independent of the adjacent vocalic context.

Although the present study is not concerned with the classificatory separability of consonantal place categories per se, the systematic differentiation between place categories and the similar ranking of place categories on a high-low F2 continuum at vowel onset and offset within each vowel category (see section 5.3.1), are consistent with Sussman's observations. Highly linear relationships between F2 values at vowel onset/offset and at the vowel midpoint for some vowels are also demonstrated in the multiple regression analyses.

The near maximal F2 context-dependency Sussman notes for bilabial consonants parallels the near maximal F2 context-dependency observed for schwa, /ɪ/ and /ʊ/ in the present data. Sussman also explains the highly linear relationship he observes for bilabials in terms of their inherent underspecification for tongue position. Because, bilabials are unspecified for tongue position, during their production, the tongue body is free to adopt the configuration required for the upcoming vowel.

The relatively stable onset frequencies Sussman reports for the approximants /w, j, l, r/ also has a parallel in the present data. /w/, /ɫ/ and, to a lesser extent /j/ are among the consonants which exert the strongest coarticulatory influence on vowels in the present study. The fact that these consonants also show the least vowel dependence in Sussman's data is further evidence of the existence of a reciprocal relationship between degree of coarticulation and strength of coarticulatory influence (Recasens, 1991; Farnetani et al., 1993).

The contextual variation observed for F1 in the present study is generally less extensive and less systematic than the variation observed for F2. This may be attributable to three factors. Firstly, because consonants are typically characterised by a lesser degree of jaw opening and hence lower F1 values than vowels, shifts in F1 as a function of consonantal context are likely to be relatively small compared with shifts in F2. The direction of influence is also always likely to be in the same direction (Stevens & House, 1963). Secondly, because variations in F1 reflect variations in the overall cross-sectional area of the vocal tract they can be less simply correlated with one or other of the phonetic dimensions of tongue and jaw height (Lindblom & Sundberg, 1971). The interactive influence of tongue, lip and jaw effects is therefore not so easily apportioned to the influence of either adjacent consonantal place or manner of articulation in an analysis of variance in which these serve as the grouping variable. Evidence in the literature also indicates that vowel F1 values vary considerably as a function of the voicing characteristic of the adjacent consonants

(Fant, 1973; Lisker, 1985). Thus, a significant proportion of F1 variance not accounted for in the present study may reflect the lack of distinction made between voiced and voiceless consonants.

A further possibility to account for the lower overall context-dependency observed for F1 compared with F2, is that F1 is subject to tighter perceptual constraints on variability. For example, Manuel (1990) suggests that, in English, less coarticulation may be tolerated along F1 than along F2 due to the fact that there are more contrastive levels of height than front/backness and hence a greater potential loss of contrast along this dimension. Greater variability for schwa along F1 than along F2 (see Figure 5.13 and Figure 7.3) may be attributed to the fact that schwa is not differentiated from any other vowel by height alone. The greater F1 compared with F2 variability observed for unaccented /ɪ/ tokens is more difficult to account for within this framework unless they are considered as fully reduced and hence non-contrastive.

It is possible that greater variability along a given dimension may be tolerated in some cases owing to the maintenance of distinctiveness by other features. For example, high F1 variability observed for /i/ is unlikely to unduly affect perceptual judgements of /i/ owing to its distinctiveness along F2. Perceptual constraints on variability may therefore be limited to those features for a given vowel which are the most perceptually salient.

Notwithstanding these points, systematic variation in vowel F1 onset, offset and midpoint value as a function of both adjacent consonantal place and manner of articulation is evident in the present data. Relatively high levels of F1 context dependency observed for schwa, /ɪ/ and /ɛ/ and the large difference in levels of context-dependency for these and other vowels across stress conditions also indicates systematic context effects. These results suggest that contextual variation in F1 warrants more attention than it has hitherto received.

7.5.2 Directionality and temporal extent of effects

The present data adds to the body of literature which reports greater carryover compared with anticipatory coarticulation (Ohde & Sharf, 1975; Bell-Berti & Harris, 1976; Gay, 1977; Fowler, 1981; Recasens, 1984; Magen, 1984; Krull, 1989; Choi, 1992). There are some differences in the directionality of effects across vowels and/or stress conditions with respect to the magnitude of effects (i.e. amount of shift in vowel formant value). However, all vowels, with the exception of /ə/ and /ɪ/, show greater carryover than anticipatory coarticulation for F2 where this is measured in terms of the degree of differentiation in F2 midpoint value as a function of individual consonantal contexts. There is less distinction between full and reduced vowels in terms of the directionality of effects along F1. However, as in the case of F2, there is generally greater carryover compared with anticipatory coarticulation.

The relatively high degree of anticipatory coarticulation along F2 observed for schwa and /ɪ/ and the relatively low degree of anticipatory coarticulation along F2 observed for the full vowels, accords with Recasens' (1984) proposal that anticipatory effects are dependent on the degree of articulatory constraint on the tongue body during production of the target vowel whereas carryover effects are largely independent of such constraints. It also lends support to the proposal that coarticulatory effects are recipricol such that segments which show the most coarticulatory resistance also exert the greatest coarticulatory influence (Recasens, 1991; Farnetani et al., 1993).

Thus, in the case of /ə/ and /ɪ/, the relatively high degree of differentiation in both offset and midpoint value as a function of following consonantal place of articulation reflects their lack of specification for tongue position. Thus, in $C^1əC^2$ and $C^1ɪC^2$ syllables, in addition to greater anticipatory consonant-to-vowel effects, this also implies an absence of carryover vowel-to-consonant effects. This is illustrated in the tighter clustering of offset values compared with onset values observed for schwa in Figures 5.1a and 5.2a compared with 5.1b and 5.2b and Figure 5.3 compared with

Figure A 2. The greater spread in C^1 locus values reflects the greater influence of the context preceding C^1 on C^1 compared with the influence of schwa on C^2.

The greater carry-over coarticulation for F2 noted in the present study may also reflect the fact that final consonants are articulated with less vocal effort than initial consonants. Choi (1992) also reports asymmetries in consonantal locus values at vowel onset and offset such that, as in the present data, generally lower values are obtained for syllable-final consonants than for the corresponding syllable-initial consonants. As Choi points out, these asymmetries may reflect aerodynamic factors. Kent & Moll (1972) suggest that there is a greater build-up in air pressure implying a greater force of articulation for initial compared with final consonants. Krull (1989) also observes less separability in the slope of regression lines for final consonants in VC syllables than for initial consonants in CV syllables, indicating a greater vowel dependence for final locus values than initial loci.

In addition to extensive anticipatory effects, schwa and /ɪ/ also both show a high degree of separability in midpoint value as a function of preceding consonantal place of articulation. A comparable magnitude and range of carry-over and anticipatory effects for schwa and /ɪ/ accords with Keating's (1988) predictions for unspecified segments. It thereby also conflicts with Browman & Goldstein's (1992) proposal that schwa is co-produced with a following full vowel. However, as Kingston (1992, same volume) demonstrates, the evidence upon which Browman & Goldstein base their conclusion is weakened if the data are re-analysed using R^2 values to evaluate the regression models rather than the standard error of estimate which they use.

The present data also demonstrate that coarticulatory influence is largely restricted to the immediately adjacent context segments (see section 5.2.2.2). The minimal effects from next adjacent context segments indicate an extensive consonant-vowel interaction such that vowel-to-vowel effects are either blocked by an intervening consonant or are manifest as vowel-to-consonant-to-vowel effects (see also section 5.4.1.4). This finding is consistent with Recasens' (1984, 1986, 1991) articulatory constraint based model of coarticulation. In a series of electropalatographic and acoustic studies, Recasens demonstrates that the magnitude, the temporal extent and the directionality of coarticulatory effects vary according to specific consonant-vowel combinations. For example, Recasens (1984) shows that degree of vowel-to-vowel coarticulation in VCV sequences is dependent upon the degree of articulatory constraint involved in the production of both the vowels and the intervening consonant. These findings accord with earlier articulatory studies by Gay (1974; 1977) which also demonstrate a strong consonantal influence on the vowel-to-vowel trajectory in VCV syllables.

Much of the research into coarticulatory effects on vowels has focussed on vowel-to-vowel coarticulation (Bell-Berti & Harris, 1976; Magen, 1984, 1989; Fowler, 1981; Manuel & Krakow, 1984; Manuel, 1990). While these studies report vowel dependent effects extending to a transconsonantal vowel, they all, with one exception (Manuel & Krakow, 1984), examine effects in VCV syllables where the intervocalic consonant is a bilabial. Because bilabials do not actively involve the tongue body in their production, they are largely transparent to vowel-to-vowel effects. Manuel & Krakow (1984) also examined vowel-to-vowel effects across an intervening /t/. However, as in the case of the other studies cited, they also looked at vowels in nonsense syllables. Vowels produced in isolated words or nonsense syllables are likely to show less variability and reduction than vowels produced in connected speech. They are, therefore, also more likely to exert a stronger coarticulatory influence. As the present results indicate, vowel-to-vowel coarticulation in connected speech data cannot be assessed independently of consonant-vowel coarticulation. The investigation of how different vowel and consonant gestures combine to influence the acoustic signal is thus an important area for future research.

7.6 Problems and benifits of a corpus-based approach

The decision to use an existing speech database as opposed to setting up a controlled experimental paradigm depends upon the nature of one's research objectives. To date, phonetic research has largely been carried out within the experimental framework. The reason for this is clear, it allows the researcher to vary the research factors of interest in a systematic and principled way and also affords control over the potential interaction of factors. A major drawback with this approach, however, is that the speech material tends to be of a limited and artificial nature. Thus, while the controlled, experimental paradigm is fundamental to phonetic research, there is also a recognised need for research into more variegated real speech data. The increasing availability of large corpera of read and/or spontaneously produced speech coupled with the technology to store and manipulate large amounts of speech data has provided new opportunities to carry out such research.

One of the principal benefits of a corpus-based approach is that it permits exploratory, observational studies of speech data which has not been collected on the basis of any *a priori* assumptions, that is, with reference to specific research questions. It thus avoids potential biases which can arise as a function of the experimental design. It thereby also provides the opportunity for investigating the generalisability of hypotheses made on the basis of highly controlled experimental data to less variationally restricted data.

Another major advantage is that access to large amounts of data permits rigorous, quantitative studies of speech phenomena. It is generally the case that the larger the sample size, the greater is the predictive power of the statistical analyses of the data, i.e. the greater the generalisability of observations based on the current speech sample to the wider population.

The relatively unrestricted nature of the speech material in a corpus-based study, however, also has disadvantages. Because it is not possible to manipulate variables in a systematic way or control for the interaction of variables, all conclusions formulated on the basis of observational data must necessarily be more tentative than those made on the basis of controlled, experimental data. For example, when assessing the effects of sentence stress on vowel duration, it is necessary to take into account that duration is also influenced by other factors such as word-length, position in word and segmental context. These factors may all vary across vowels and stress conditions. Thus, the possibility remains that the observed differences across vowels with respect to the relative decrease in duration as a function of sentence stress reflects differences in the frequency distribution of these factors rather than inherent properties of the vowel. If the speech sample is large enough, it may be possible to treat these factors as random variables. However, even with very large speech samples, with unrestricted data there is always the possibility that the results to some extent reflect the influence of unidentified factors. In addition to creating spurious relationships, unidentified factors may also serve to obscure or surpress real relationships which exist between variables.

Another potential drawback in using an existing database concerns the level of confidence one may have in the data. In working with a database that has been constructed by someone else, it is essential to be familiar with the labelling criteria used and to be satisfied that these have been applied consistently. The extent to which the researcher feels it necessary to check the data will depend on factors such as where and when, for whom and for what purpose it was collected. For example, inhouse databases may be less reliable than databases created for use in the public domain. Depending on the size of the corpus and the nature of any amendments they may wish to make, the researcher may thus find themselves with a considerable amount of preliminary work. This possibility should be taken into account when deciding on the research approach.

Given that the researcher has to work within the design constraints of the database, they may also find themselves limited by the amount and nature of the available speech material. For example, syllable boundaries may not be marked or there may be insufficient examples of a particular segment or segmental sequence of interest. However, this should not be a problem if the research objectives are clearly defined in advance.

In sum, a corpus-based approach is attractive principally because it offers the opportunity for examining relatively large amounts of meaningful connected speech data. The study of connected speech data is important not least because reduction phenomena are conditioned by prosodic, semantic and pragmatic in addition to segmental factors and it is difficult if not impossible to simulate the interaction of these variable in highly controlled experimental data. There is also a need to define the limits of phonetic variability in real speech and to test the generalisability of hypothesis made on the basis of more restricted data. However, the lack of experimental control implies a certain degree caution in the interpretation of results.

By their very nature, the corpus-based approach and the experimental approach afford different research possibilities. They, therefore, both have an important role to play in phonetic research, the findings from investigations using one approach ideally serving to fuel further investigation using the other approach.

7.7 Future research

The present study describes the pattern of variability that occurs in the vowel system of one male speaker of one accent of British English and for one particular speech mode. In order to generalise from the present results, it is necessary to examine patterns of variability for other speakers of the same and different accent systems and for other speaking styles. Speakers are known to employ different coarticulatory strategies (Nolan, 1985). Differences in the size of the vowel space have also been observed for speakers of the same accent system (Van Bergem, 1993). Gender and age- related differences in pronunciation have also been documented (Henton, 1983). Pronunciation is also known to vary considerably across different speaking styles (Koopmans van Beinum, 1980; Lindblom & Moon, 1988; Krull, 1989; Harmegnies & Poch-Olive, 1992). Thus, different speech samples are likely to show considerable variation with respect to the overall level of phonetic variability they exhibit. Furthermore, depending on the accent system (or language), differences between vowels in terms of inherent variability might also be expected on the basis of phonological and perceptual constraints (Manuel, 1990; Lindblom & Engstrand, 1989; Lindblom, 1990).

In addressing the question of inherent vowel variability, it is also important to establish how far the observed patterns of acoustic variability reflect the corresponding pattern of articulatory variability. Studies which examine articulatory and acoustic data in parallel should therefore have a more prominent role in future research.

In order to increase the predictive power of the window model of coarticulation, further work is also required to establish the ranges of variability for different segments along different phonetic dimensions and to derive the appropriate interpolation functions. As Keating (1990) herself concedes, much of the fine working detail for the window model remains to be developed. However, a framework in which gradient, coarticulatory phenomena can be adequately described becomes increasingly important as research progresses further into the domain of natural speech data. The window model, potentially, offers such a framework.

7.8 General summary and conclusions

This thesis had three related objectives. The first of these was to assess the magnitude and patterns of contextual variability displayed by schwa, the central or 'reduced' vowel in English, with reference to the question of whether or not schwa can be associated with an independent phonetic target. Reports from two earlier studies (Browman & Goldstein, 1992; Van Bergem, 1994) present conflicting conclusions with respect to this question. These studies were also based on nonsense data although Van Bergem tested his model on a small subset of meaningful data. In the present investigation, data is examined for schwa tokens produced in a more comprehensive range of contexts and in meaningful connected speech. The present study also provides a comparative measure of schwa's variability and hence a more reliable evaluation of schwa's targeted/targetless status.

The second objective was to compare patterns of contextual variability for the full vowels in order to determine whether some vowels are inherently more susceptible to coarticulatory effects than others (e.g. lax (short) vowels compared with tense (long) vowels or back vowels compared with front vowels). This question has received relatively little attention in the literature since most studies which examine coarticulatory effects tend to use a restricted number of vowels. The present study presents a quantitative evaluation of the relative contextual variability shown by the twelve monophthongs for one male speaker of SBS.

The third objective was to assess the role of sentential stress in conditioning vowel quality. Lindblom's (1963) undershoot hypothesis proposes that reduction in vowel quality as a function of reduced stress reflects increased contextual assimilation. This idea contrasts with the traditional concept of vowel reduction as an independent process of articulatory and acoustic centralisation. Subsequent research provides some evidence that (lexically and/or sententially) unstressed vowels are more sensitive to coarticulatory effects than their stressed counterparts. However, for the most part, this has also been based on nonsense data or isolated words produced in semantically empty carrier phrases. A principal motivation in the present study, therefore, was to investigate the effects of stress in more natural speech data.

The unifying principle with respect to the three objectives, was to obtain a better understanding of the nature of vowel reduction in English. Specifically, the data was examined for evidence to support an account of vowel reduction in terms of phonetic underspecification.

The results for the pooled data analysis support the hypothesis that schwa is completely unspecified for tongue position. They also indicate that schwa is highly unspecified for jaw position. In this respect, the present results for English accord with Van Bergem's (1994) observations for Dutch schwa. Significant differences between the full vowels with respect to degree of context-sensitivity is also consistent with the proposal that segments may show varying degrees of specification along a given dimension(s) (Keating, 1990).

It is proposed that the observed hierarchy of vowel robustness, as measured in terms of overall degree of context-dependency, reflects quantal, acoustic properties of vowels as described by Stevens (1989) rather than inherent duration. The present acoustic data are also consistent with expectations regarding articulatory and perceptual constraints on variability. It is suggested that the lesser overall degree of context-dependency noted for the more peripheral vowels /i, a, ɑ, ɒ, ɔ/ compared with the more central vowels /ɪ, ɛ, ʌ, ʊ, ɜ/ and, for the present speaker, /u/, reflects a lesser overall degree of gestural compatiblity with context and consequently a lesser overall degree of contextual assimilation. Greater articulatory and/or acoustic stability for the peripheral vowels is also consistent with the idea that these vowels serve to calibrate the vowel space for a given speaker.

In general, the effects of sentence stress were found to be additive. Variations in sentence stress caused variations in the relative magnitude of the overall coarticulatory effect rather than in the nature or directionality of the coarticulatory influence. For the most part, the same hierarchy of vowel robustness was evident within each accent condition. In most cases, the sententially stressed/unstressed comparison of the data showed a significant increase in amount of context-dependency for unaccented vowel tokens relative to their accented counterparts. The vowels which did not show a significant stress effect or which showed the least difference as a function of sentence stress were either those vowels which showed the least context-dependency overall or the vowels which showed the greatest overall context-dependency.

Evidence that schwa is targetless and can therefore occupy almost any position in the vowel space depending on context indicates that vowel reduction is a matter of contextual assimilation rather than centralisation. Greater context-sensitivity for unaccented relative to accented vowels is also consistent with an account of vowel reduction in terms of contextual assimilation. The overall reduction in acoustic contrast observed for unaccented vowel tokens relative to accented tokens may be attributed to the averaging of individual coarticulatory shifts.

These results are also consistent with the view that vowel reduction represents a means of economising on articulatory effort. Schwa the endpoint of the reduction process arguably represents minimal articulatory effort insofar as it represents the straight-line interpolation between context segments and hence minimal resistance to coarticulatory effects. Shorter durations, greater context-dependency and, in the case of the peripheral vowels, less extreme formant values for unaccented compared with accented tokens is also consistent with studies in the literature which report reduced muscle activity (Gay & Hirose, 1974; Gay, 1978; Tuller, Kelso & Harris, 1982) and/or reduced articulatory displacement (Delattre, 1969; Kuehn & Moll, 1976; Summers, 1987; Engstrand, 1988; Flege, 1988; Farnetani & Faber, 1992) for (lexically and/or sententially) unstressed vowels. Generally greater contextual variability for lax vowels compared with tense vowels is also consistent with the hypothesis that lax vowels are more "cost-effective" (see section 2.7).

On the basis of these results, it is proposed that vowel reduction may be accounted for in terms of phonetic underspecification operating along two parallel continuua: a peripheral-non-peripheral continuum and a stressed-unstressed continuum. Although it is untestable in the present study, it is proposed that these continuua both reflect differences between vowels with respect to the degree of articulatory effort involved in their production. However, whereas the peripheral-non-peripheral continuum reflects inherent differences in vowel quality, the stressed-unstressed continuum is assumed to be part of a larger independent continuum of hyper-hypospeech which is style-conditioned. A separate specification for unstressed compared with stressed vowels is nevertheless advocated on the grounds that the reduction in quality reflects a 'planned' reduction in vocal effort. It is not simply the automatic consequence of a reduction in duration and of mechano-inertial effects. However, it is acknowledged that, in practise, it may be difficult to distinguish between the two effects.

A principal advantage of a phonetic underspecification account of English vowel reduction is that phonetic and phonological vowel reduction may be accounted for by the same mechanism. Phonetic vowel reduction is a continuous coarticulatory effect which applies in varying degrees to all vowels produced in context. It is proposed that non-obligatory phonological reduction, although lexically discrete, is also gradual at the phonetic level (see section 7. 4). Degree of reduction depends on the speaker's ambition to produce a full vowel and is conditioned by factors such as stress, word class, word frequency as well as stylistic considerations. Obligatory phonological reduction is both lexically and phonetically discrete. However, it is posited that the schwa which appears the citation form of words (the endpoint of obligatory phonological reduction) has the same phonetic characteristics as schwa which represents a fully reduced full vowel (the endpoint of phonetic and non-obligatory phonological reduction), that is, it is an empty time slot which is interpolated through.

In sum, it is proposed in this thesis that vowel reduction reflects phonetic underspecification. The corner vowels /i, a, ɑ/ and the back, rounded vowels /ɒ/ and /ɔ/ are the most narrowly specified vowels and the most resistant to coarticulatory effects. The more central vowels /ɪ, ɛ, ʌ, ʊ, u, ɜ/ are less narrowly specified and accordingly show greater contextual assimilation. Sententially stressed vowels are more narrowly specified than sententially unstressed vowels and are also less susceptible to coarticulatory effects. Schwa which shows the shortest inherent duration and which always occurs in unstressed position, is completely unspecified for tongue position and accordingly shows maximal context-dependency.

Figure 7.1: Tense (long) vowels: pooled across all contexts and accent conditions.

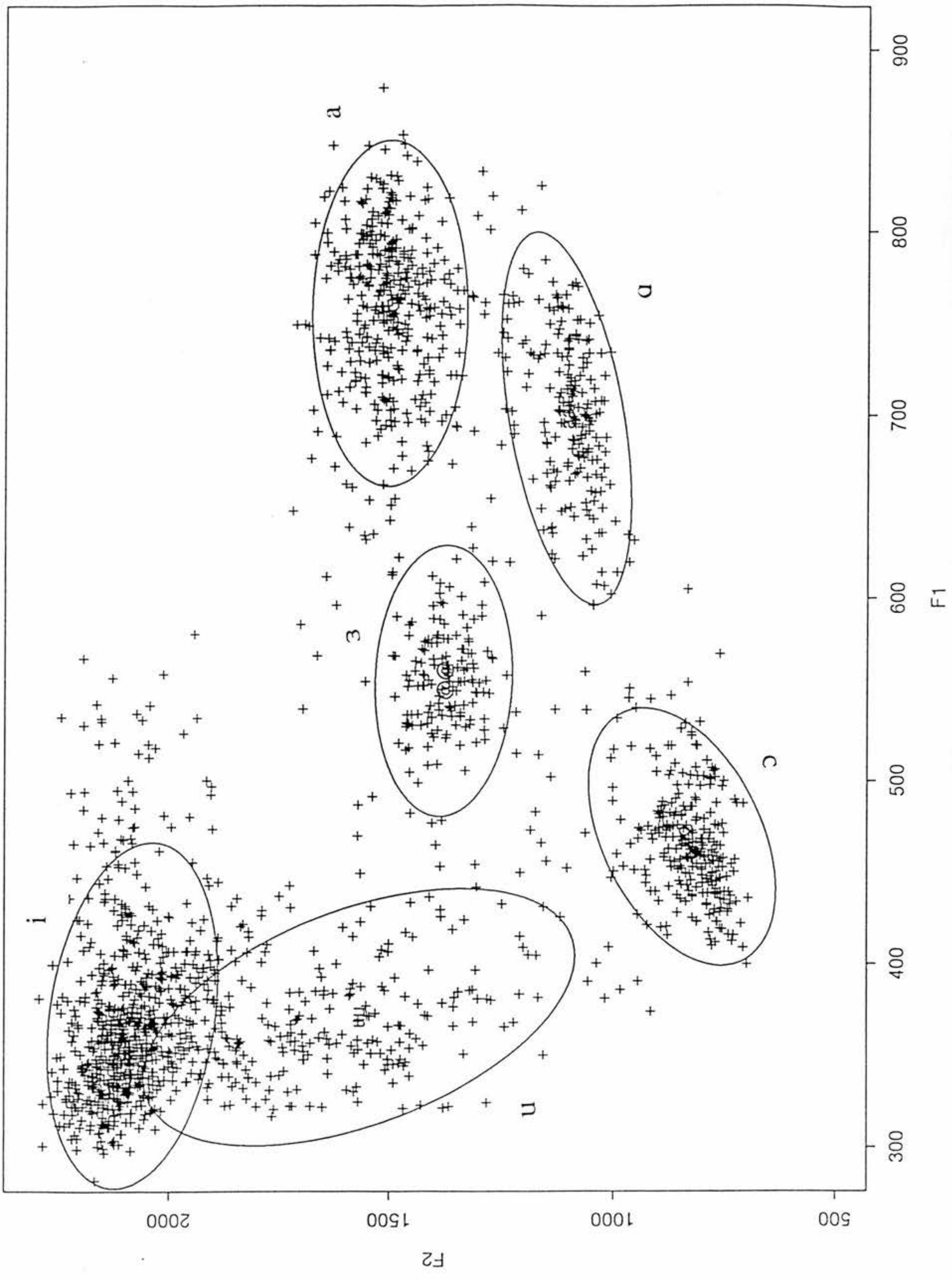


Figure 7.2: Lax (short) vowels: pooled across all contexts and accent conditions.

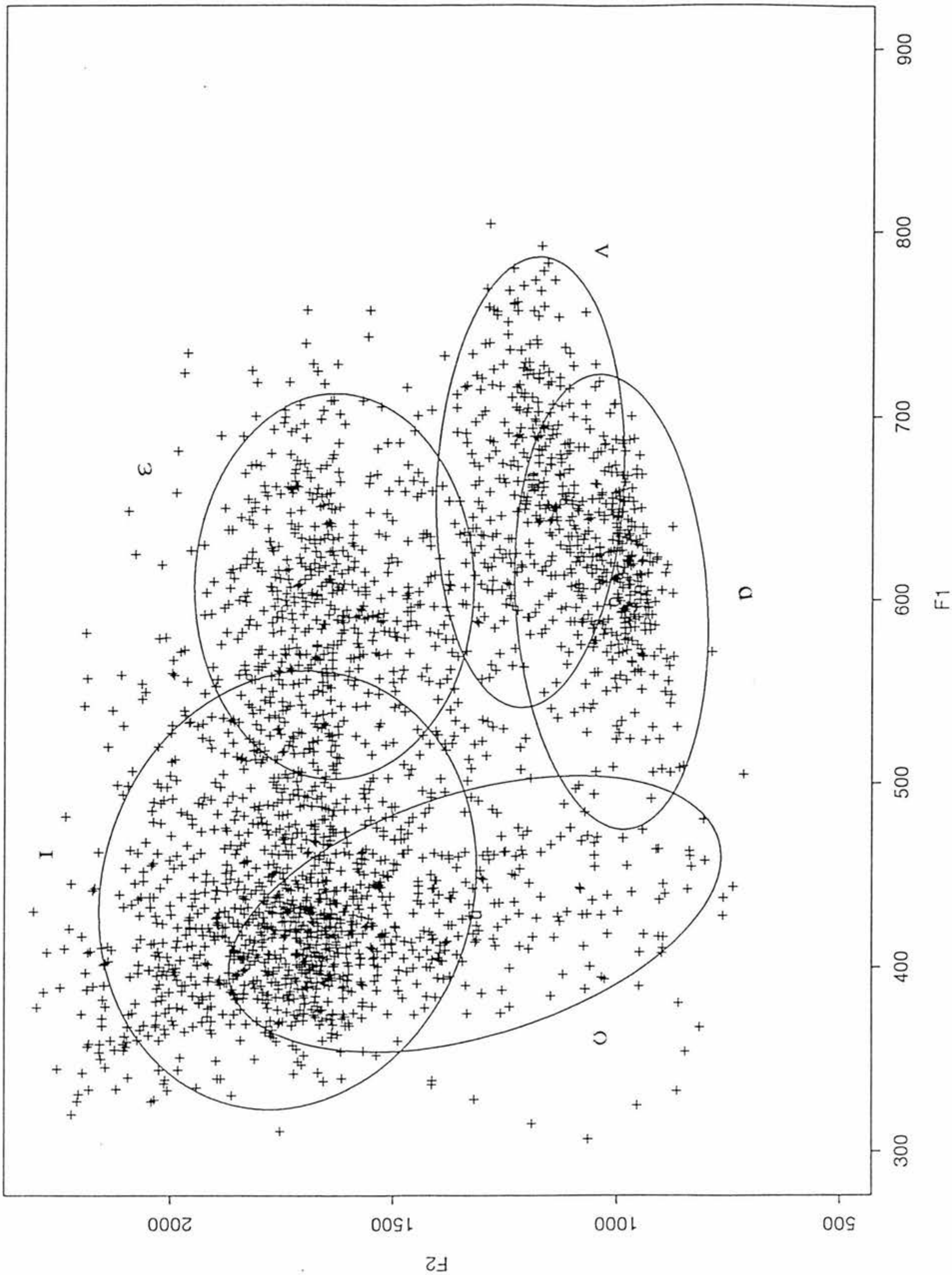
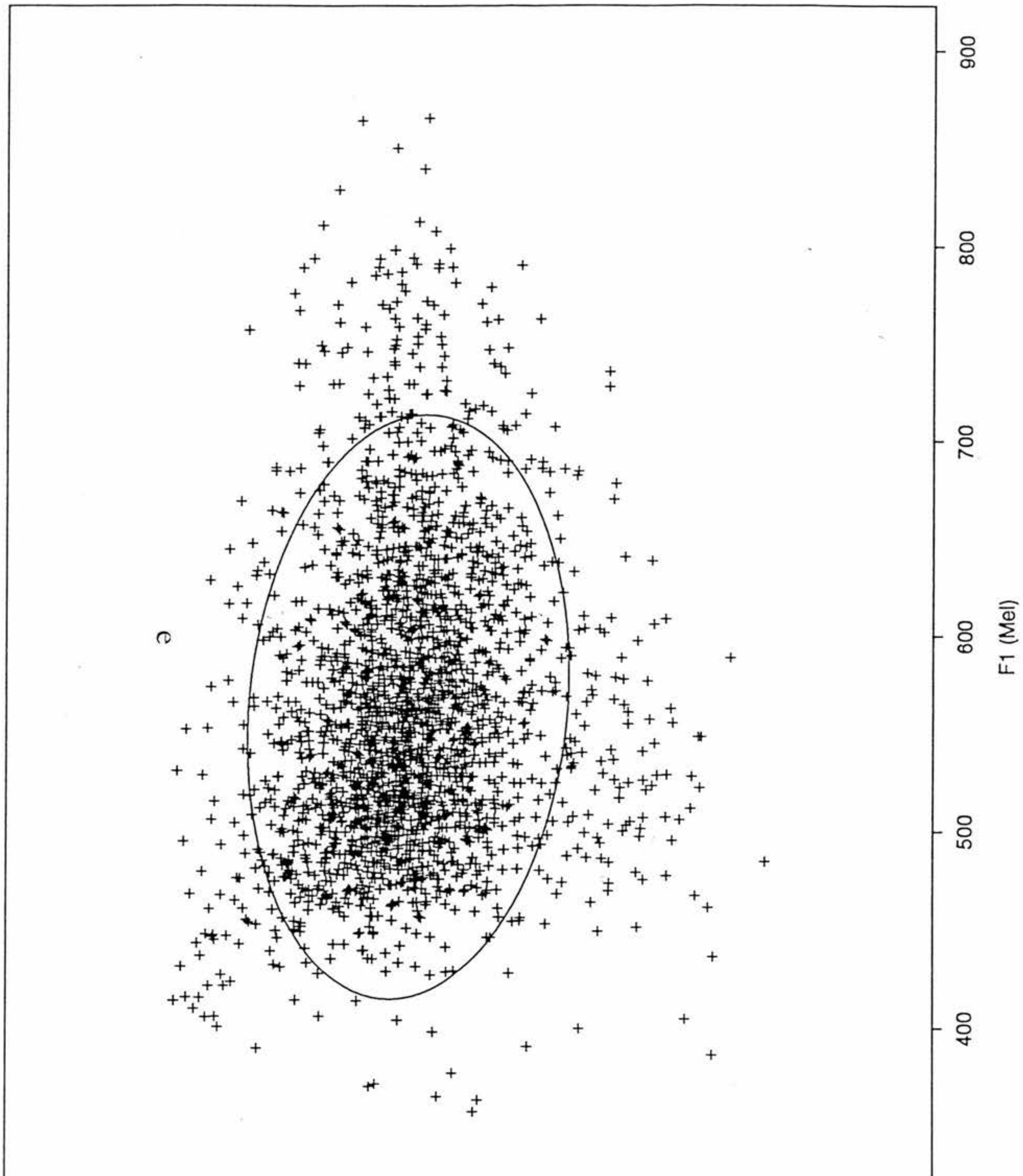


Figure 7.3: Schwa tokens: pooled across all contexts.



Appendix A

Figure A 1.a:l: Proportional distribution of preceding and following contexts for each vowel.

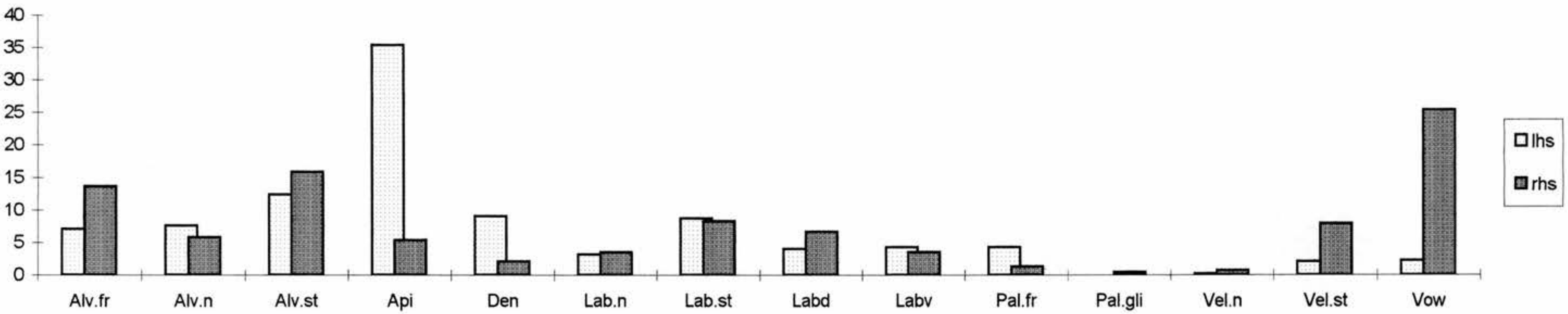
Figure A 2.a:f: Mean second formant trajectories for schwa as a function of following consonant place of articulation.

Table A 1.a:k: F2 mean onset, midpoint and offset values and measures of deviation from linearity for the full vowels.

Table A 2.a:k: F1 mean onset, midpoint and offset values and measures of deviation from linearity for the full vowels.

Figure A 1: Proportional distribution of preceding (lhs) and following (rhs) contexts

A 1a: /i/



A 1b: /ɪ/

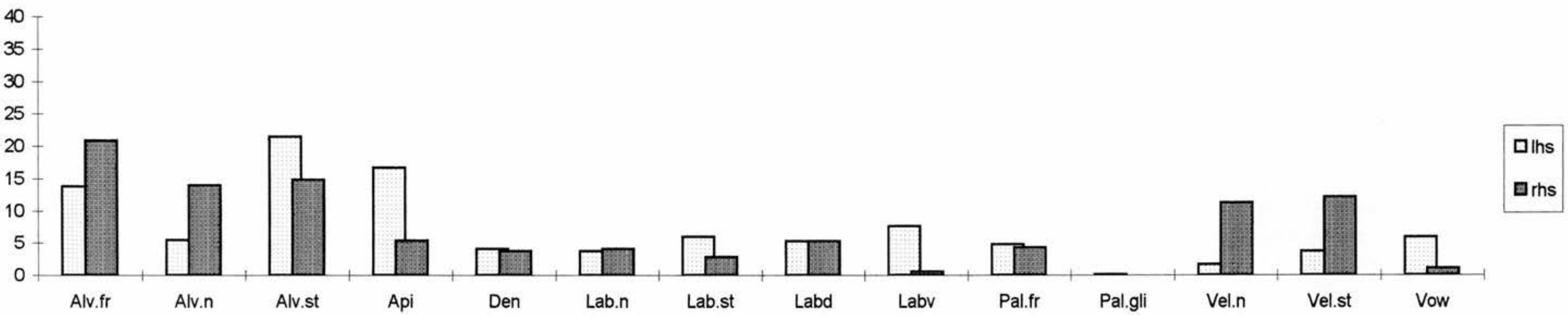
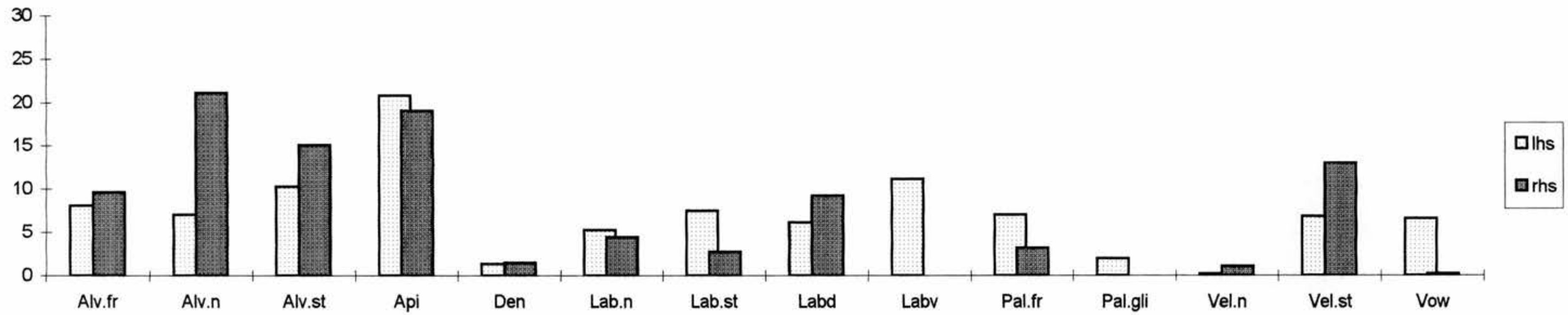


Figure A 1: Proportional distribution of preceding (lhs) and following (rhs) contexts (continued)

A 1c: /ɛ/



A 1d: /a/

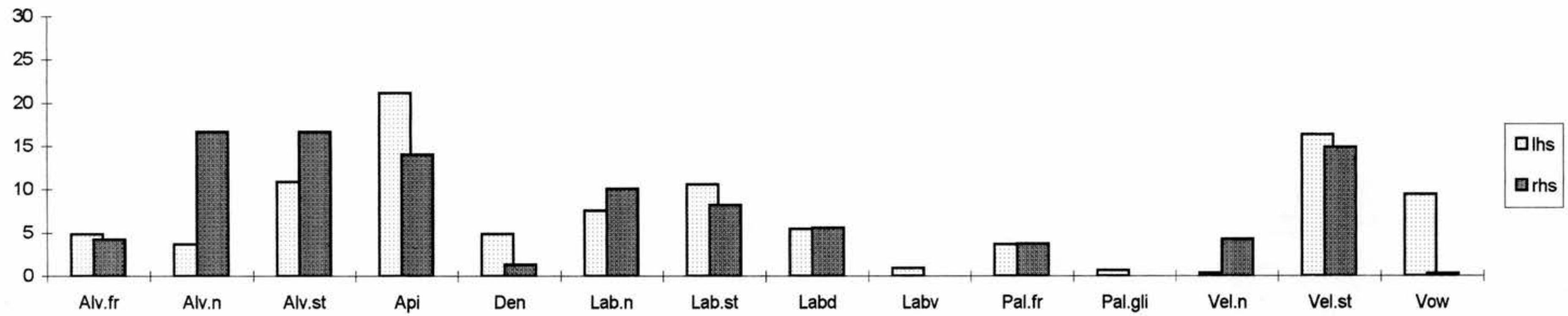
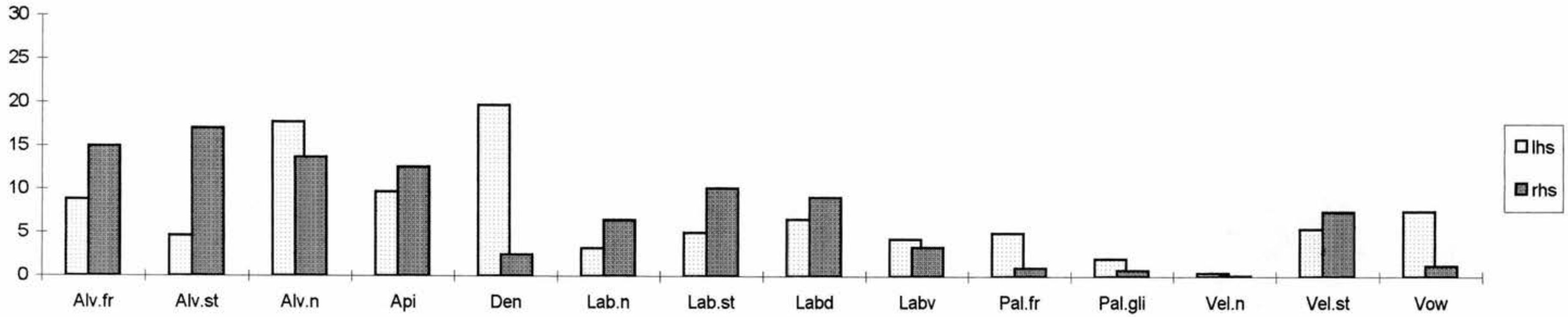


Figure A 1: Proportional distribution of preceding (lhs) and following (rhs) contexts (continued)

A 1e: /ə/



A 1f: /ɜ/

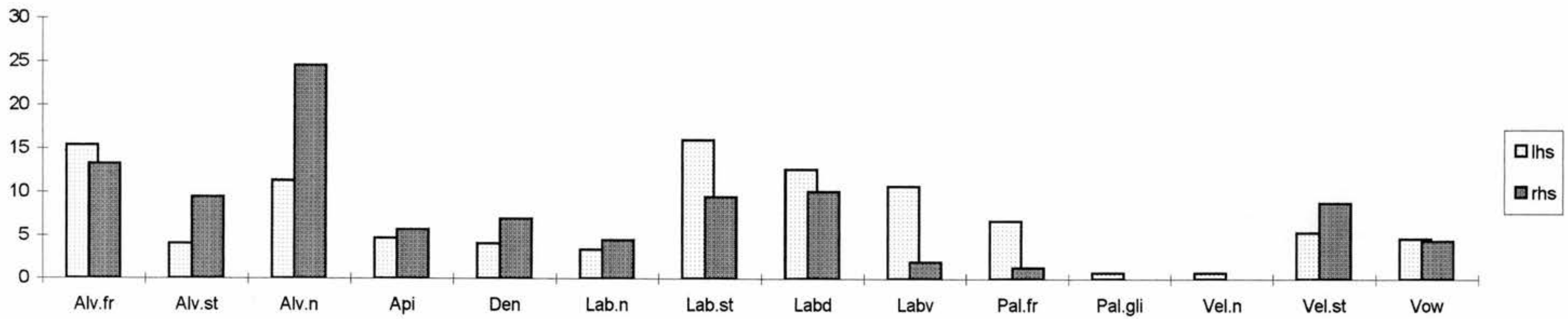
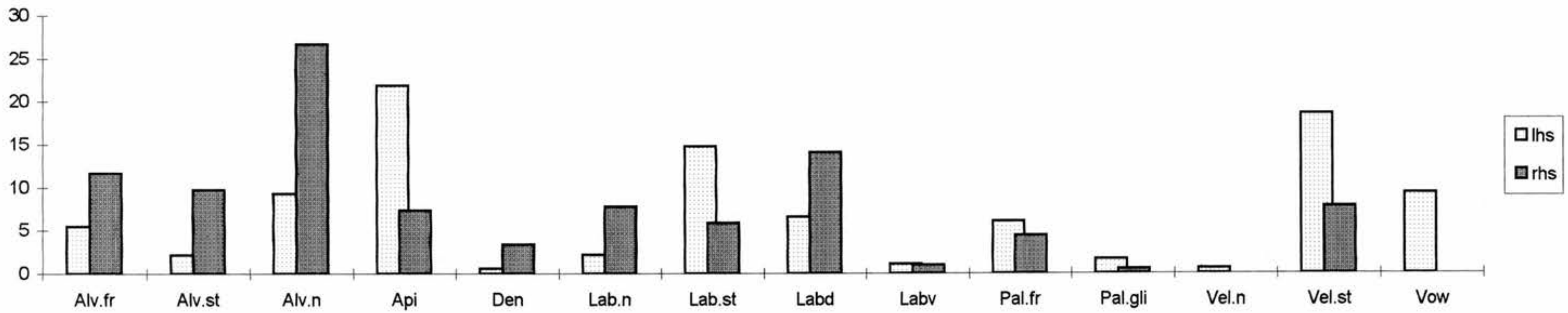


Figure A 1: Proportional distribution of preceding (lhs) and following (rhs) contexts (continued)

A 1g: /ɑ/



A 1h: /ʌ/

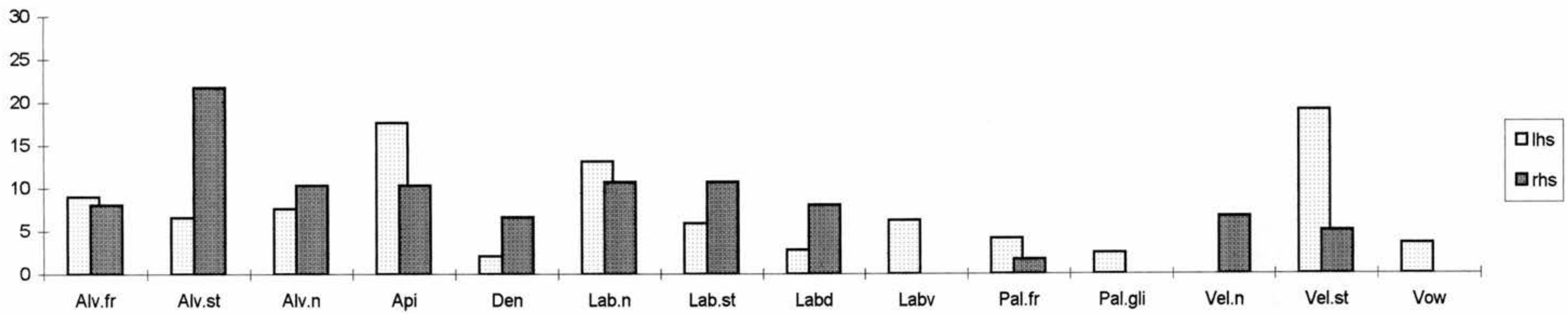
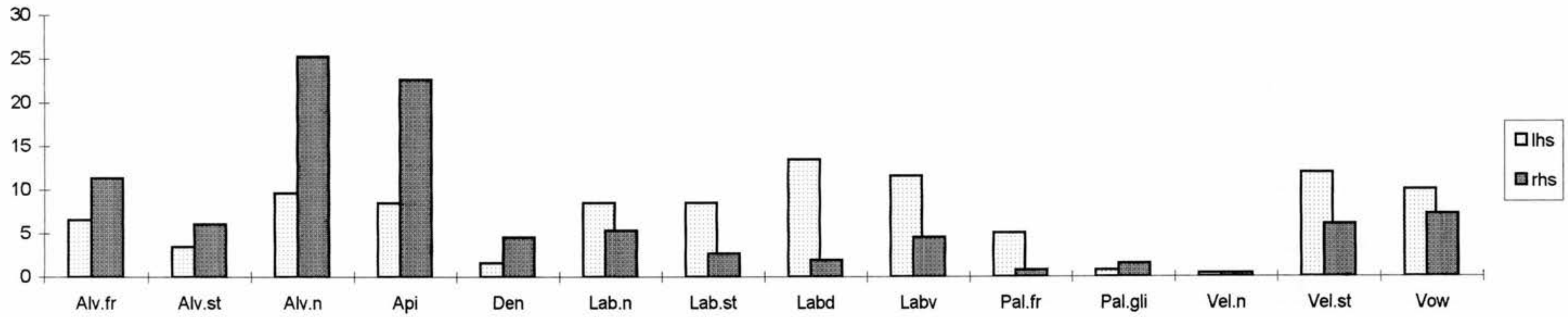


Figure A 1: Proportional distribution of preceding (lhs) and following (rhs) contexts (continued)

A 1i: /ɔ/



A 1j: /ɒ/

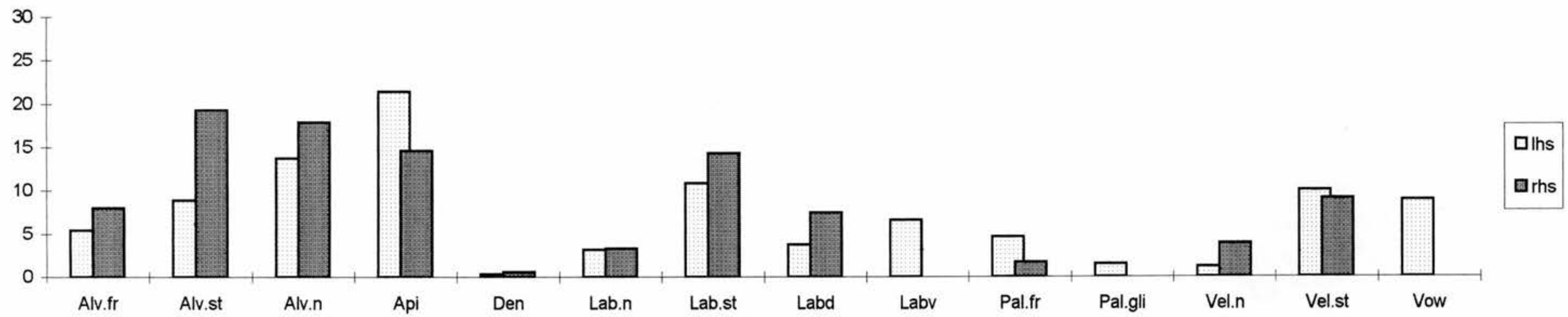
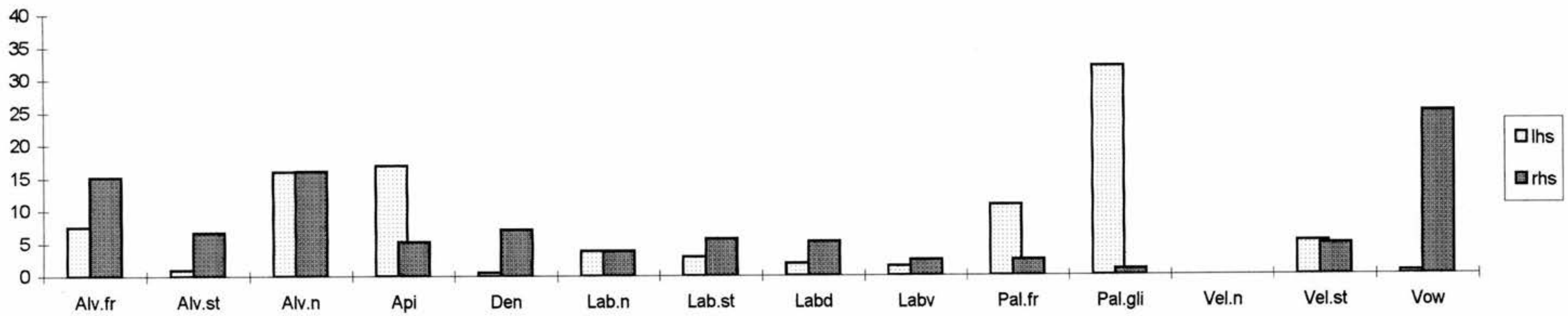


Figure A 1: Proportional distribution of preceding (lhs) and following (rhs) contexts (continued)

A 1k: /u/



A 1l: /ʊ/

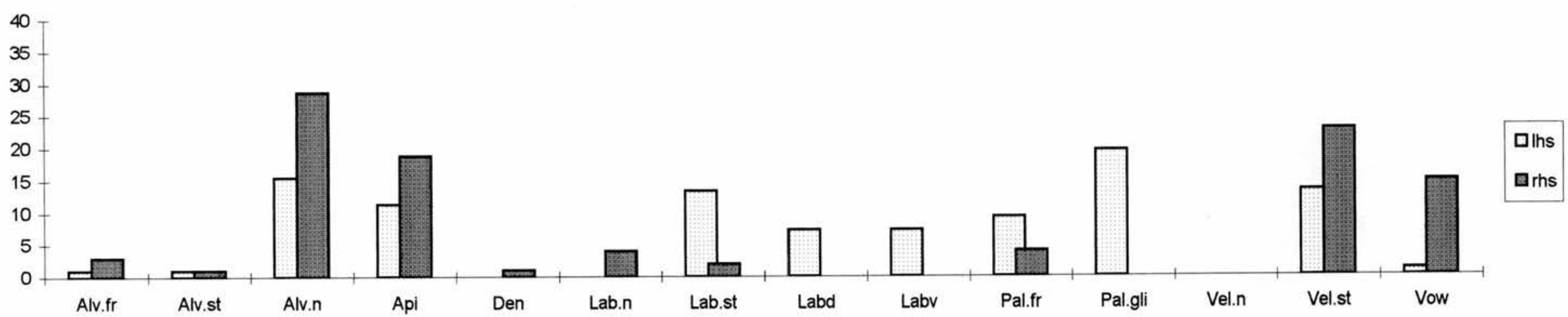
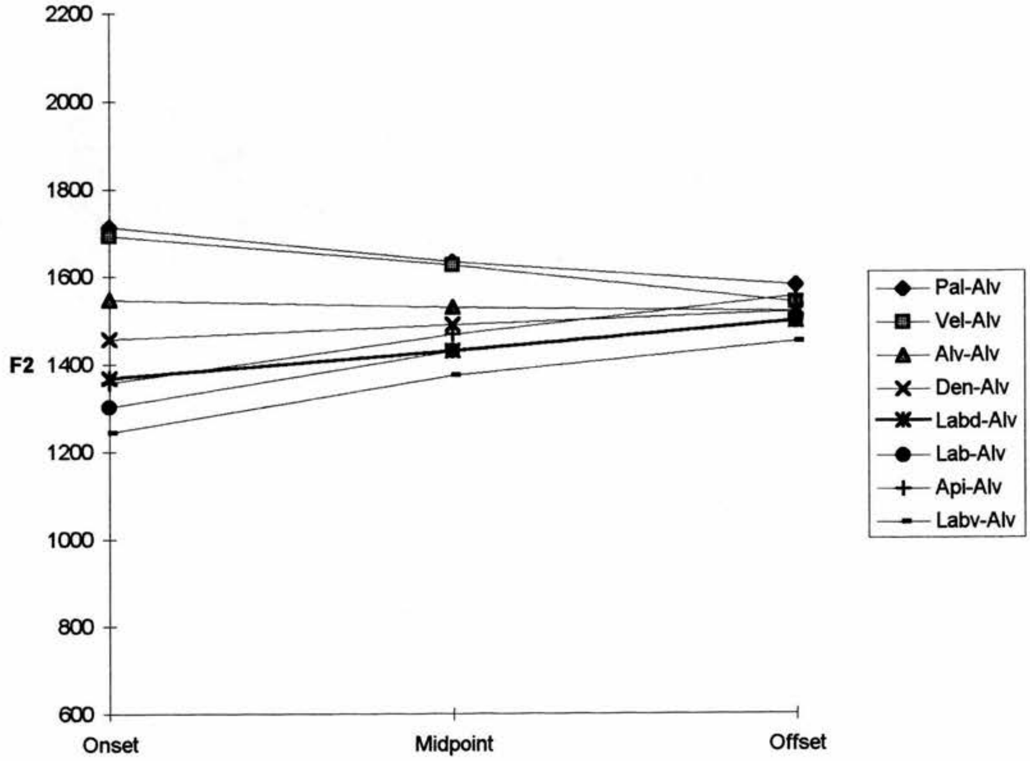


Table A 2: Mean second formant trajectories for schwa as a function of following consonant place of articulation

A 2a: alveolar



A 2b: dental

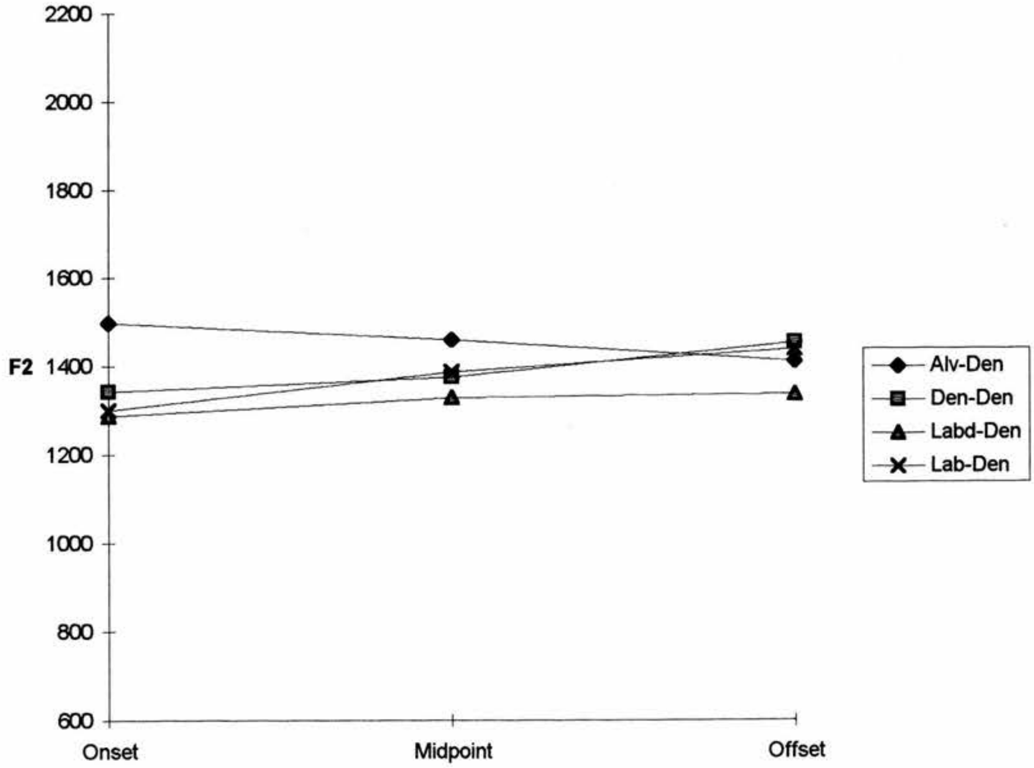
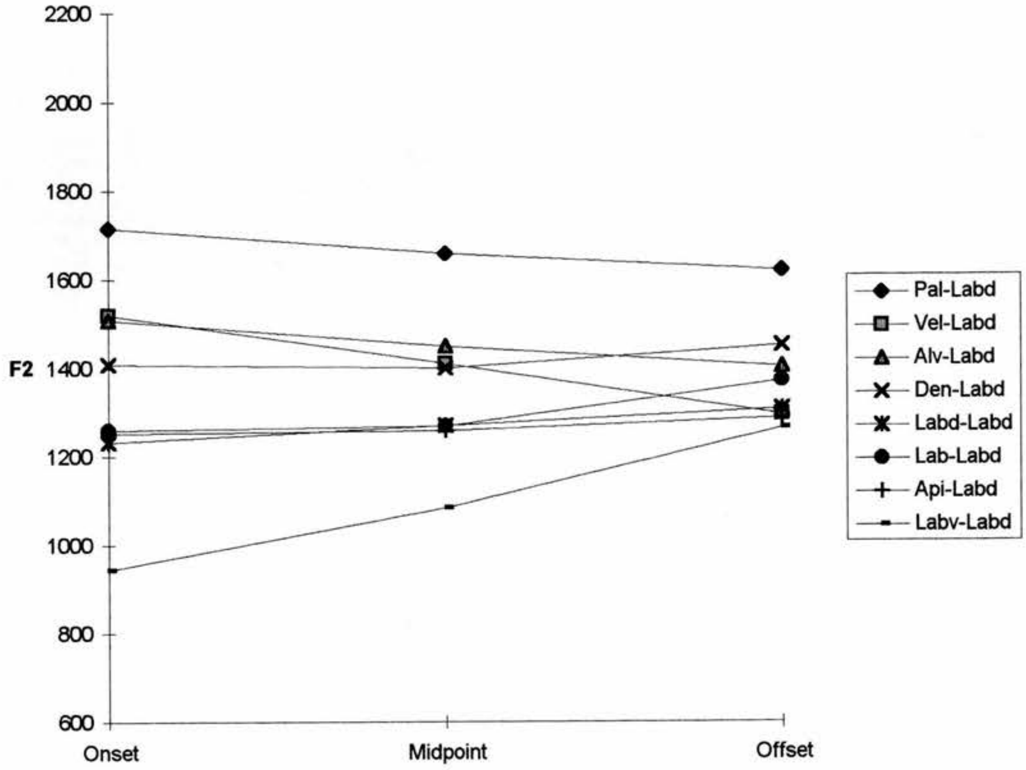


Table A 2: Mean second formant trajectories for schwa as a function of following consonant place of articulation (continued)

A 2c: labio-dental



A 2d: labial

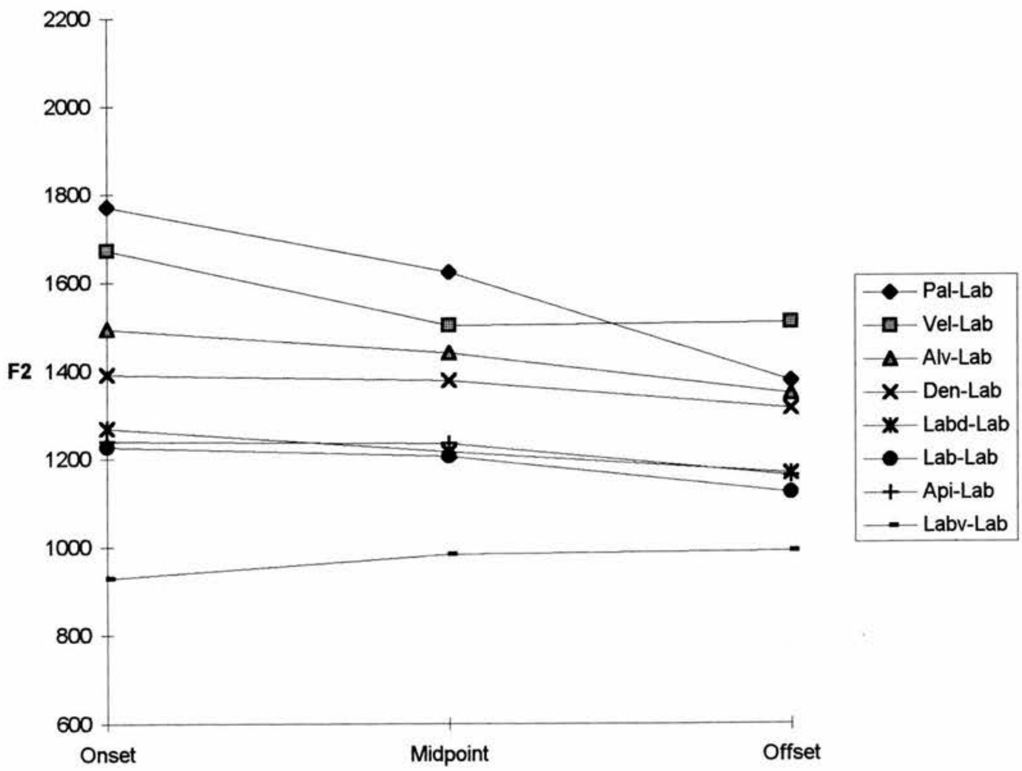
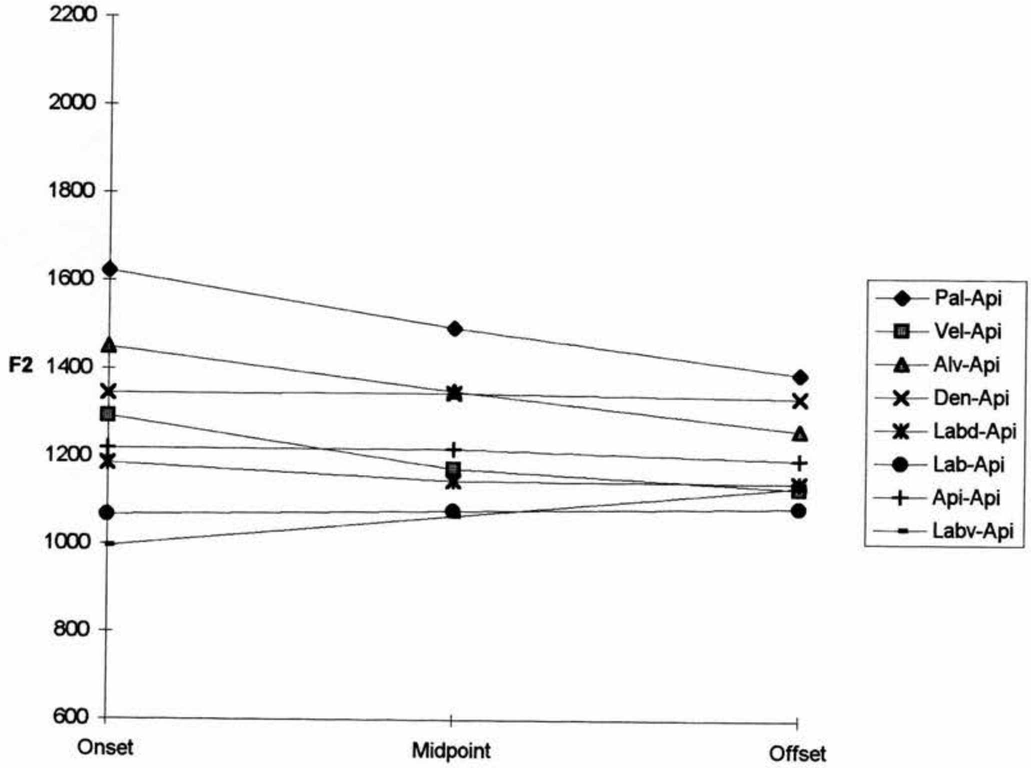


Table A 2: Mean second formant trajectories for schwa as a function of following consonant place of articulation (continued)

A 2.e: apical



A 2.f: labio-velar

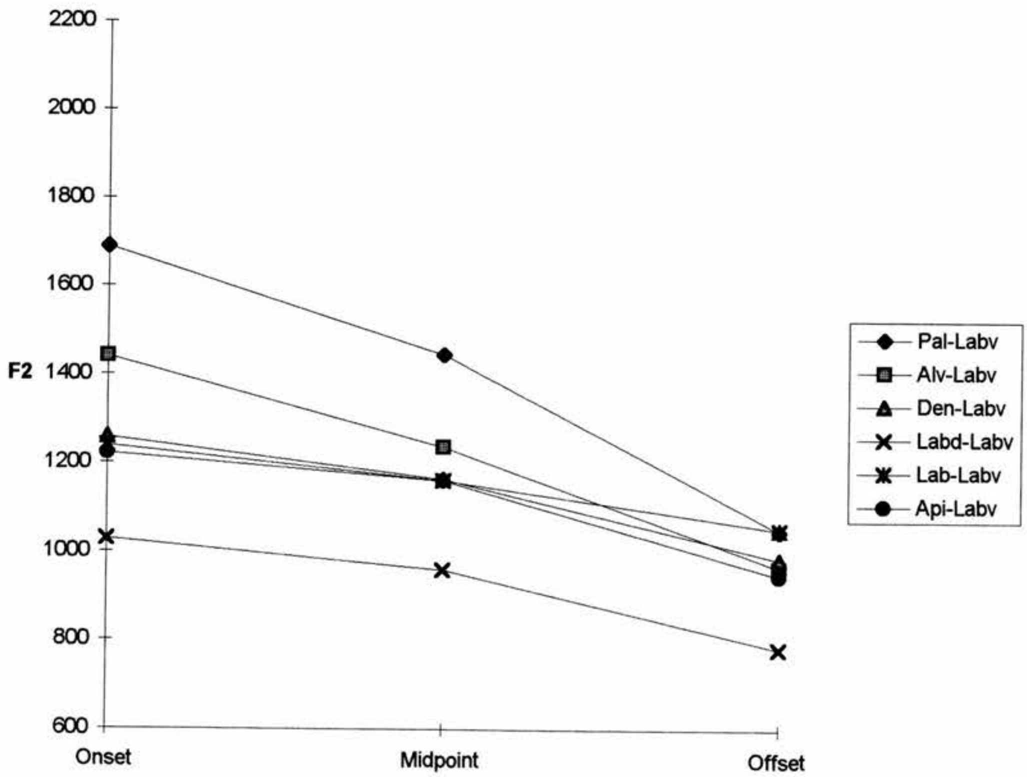


Table A 1.a: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /i/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	2001	2106	2111	67	2040	2084	2003	83	*	*	*	*
VEL	*	*	*	*	2102	2130	2205	31	2131	2162	2028	110	*	*	*	*
ALV	*	*	*	*	1980	2131	2186	64	1947	2133	2019	200	2041	2172	1994	206
DEN	*	*	*	*	*	*	*	*	1867	2063	2016	162	*	*	*	*
LABD	*	*	*	*	*	*	*	*	1833	2114	2028	245	*	*	*	*
LAB	*	*	*	*	1817	2062	2187	80	1874	2120	2072	196	*	*	*	*
API	1716	2000	2080	136	1772	2062	2164	125	1795	2063	1983	232	1928	2049	1884	191
LABV	*	*	*	*	1505	2054	2166	291	1610	2052	1865	419	*	*	*	*

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	2007	2004	1699	201	*	*	*	*	*	*	*	*
VEL	*	*	*	*	2125	2138	1956	130	*	*	*	*	*	*	*	*
ALV	1955	2141	2008	213	1949	2140	1890	294	1926	2061	1832	243	1971	2148	1738	391
DEN	1984	2108	2108	93	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	*	*	*	*	1829	1923	1664	235	*	*	*	*
LAB	1966	2087	2095	75	1943	2080	1851	244	1832	1986	1681	191	1700	2074	1193	529
API	1747	2054	1934	285	1804	2061	1881	291	1882	2012	1805	225	1727	2079	1638	529
LABV	*	*	*	*	1691	1945	1513	457	1464	1885	1740	377	1625	1810	1512	322

Table A 1.b: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɪ/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	1840	1850	1715	97	1836	1715	1610	11	*	*	*	*
VEL	*	*	*	*	2062	2058	1708	231	1931	1928	1723	135	*	*	*	*
ALV	1748	1832	1876	27	1817	1960	1865	159	1682	1683	1630	36	*	*	*	*
DEN	*	*	*	*	1682	1919	1919	143	1612	1673	1633	21	*	*	*	*
LABD	1595	1800	1902	69	1659	1842	1788	158	1559	1653	1635	67	*	*	*	*
LAB	1522	1674	1778	32	1693	1879	1934	87	1539	1707	1691	123	*	*	*	*
API	1529	1693	1813	29	1570	1747	1727	131	1414	1575	1624	75	1408	1618	1484	229
LABV	1208	1466	1722	1	1529	1758	1877	73	1305	1594	1672	141	1313	1445	1509	45

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	1842	1755	1496	115	1672	1516	1156	136	1804	1699	1585	6
VEL	1971	1891	1636	117	*	*	*	*	1793	1642	1341	100	*	*	*	*
ALV	1650	1606	1511	34	1658	1654	1496	103	1596	1519	1332	73	*	*	*	*
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	*	*	*	*	1595	1531	1320	98	*	*	*	*
LAB	1304	1302	1442	95	1643	1688	1511	148	1510	1530	1326	149	*	*	*	*
API	1503	1659	1558	71	1452	1503	1379	117	*	*	*	*	*	*	*	*
LABV	*	*	*	*	1381	1567	1386	245	1104	1206	1249	39	*	*	*	*

Table A 1.c: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɛ/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	1936	1748	1829	179	1751	1713	1872	131	1869	1733	1651	36	*	*	*	*
VEL	1884	1795	1870	109	*	*	*	*	1963	1792	1644	15	*	*	*	*
ALV	1575	1825	1873	135	1597	1760	1799	83	1606	1688	1631	93	*	*	*	*
DEN	*	*	*	*	*	*	*	*	1635	1711	1630	105	*	*	*	*
LABD	1429	1659	1772	78	1326	1686	1925	81	1452	1675	1690	139	*	*	*	*
LAB	1440	1663	1819	45	1533	1718	1821	55	1532	1726	1699	147	*	*	*	*
API	1435	1633	1771	40	1405	1686	1869	65	1449	1654	1654	137	*	*	*	*
LABV	*	*	*	*	*	*	*	*	1232	1603	1603	122	1008	1308	1361	165

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	1783	1657	1577	31	*	*	*	*	1764	1568	1275	65	*	*	*	*
VEL	*	*	*	*	1809	1741	1471	135	*	*	*	*	*	*	*	*
ALV	1611	1634	1335	215	1658	1634	1335	183	1533	1425	1209	72	*	*	*	*
DEN	*	*	*	*	*	*	*	*	1498	1434	1401	21	*	*	*	*
LABD	*	*	*	*	*	*	*	*	1396	1330	1271	5	*	*	*	*
LAB	*	*	*	*	1330	1544	1349	273	1375	1438	1188	209	*	*	*	*
API	1410	1606	1536	177	1411	1533	1380	183	1320	1446	1319	169	*	*	*	*
LABV	*	*	*	*	1193	1625	1460	398	1090	1327	1178	257	*	*	*	*

Table A 1.d: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /a/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	1808	1482	1984	552	1786	1511	1514	185	*	*	*	*
VEL	1701	1540	1702	215	1801	1671	1591	33	1732	1575	1508	60	1778	1516	1443	126
ALV	*	*	*	*	1532	1532	1581	33	1538	1522	1502	3	*	*	*	*
DEN	*	*	*	*	1376	1589	1651	101	1372	1433	1489	3	*	*	*	*
LABD	*	*	*	*	1333	1497	1825	109	1379	1458	1512	17	*	*	*	*
LAB	1246	1471	1710	9	1337	1522	1817	73	1335	1529	1544	119	*	*	*	*
API	1271	1487	1702	1	1277	1539	1723	52	1287	1482	1538	93	*	*	*	*
LABV	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	1717	1532	1333	9	1622	1428	1314	53	*	*	*	*
VEL	*	*	*	*	1732	1509	1277	6	1774	1501	1303	50	*	*	*	*
ALV	*	*	*	*	1478	1495	1275	158	1483	1494	1204	201	*	*	*	*
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	1371	1440	1091	279	1270	1302	1263	47	*	*	*	*
LAB	*	*	*	*	1517	1487	1362	63	1418	1397	1289	58	*	*	*	*
API	1368	1458	1512	24	1337	1497	1346	207	1309	1309	1255	77	*	*	*	*
LABV	*	*	*	*	1284	1501	1324	263	*	*	*	*	*	*	*	*

Table A 1.e: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɜ/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	*	*	*	*	1732	1434	1524	259	*	*	*	*
VEL	*	*	*	*	*	*	*	*	1618	1424	1525	197	*	*	*	*
ALV	*	*	*	*	1435	1348	1502	161	1507	1451	1565	113	*	*	*	*
DEN	*	*	*	*	*	*	*	*	1389	1430	1506	23	*	*	*	*
LABD	*	*	*	*	*	*	*	*	1233	1361	1535	31	1272	1302	1269	42
LAB	*	*	*	*	1155	1326	1407	60	1220	1364	1608	67	1253	1356	1447	8
API	*	*	*	*	*	*	*	*	1215	1393	1601	20	*	*	*	*
LABV	*	*	*	*	*	*	*	*	989	1297	1476	86	1074	1530	1488	332

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
VEL	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
ALV	1427	1404	1349	21	1437	1406	1256	79	*	*	*	*	*	*	*	*
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	1250	1322	1125	179	*	*	*	*	*	*	*	*
LAB	1309	1345	1344	25	1286	1310	1174	107	*	*	*	*	*	*	*	*
API	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABV	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Table A 1.f: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɑ/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	1518	1074	1605	650	*	*	*	*	1564	1110	1437	521	*	*	*	*
VEL	*	*	*	*	1460	1109	1401	429	1415	1100	1327	361	*	*	*	*
ALV	1353	1111	1457	392	1383	1071	1424	443	*	*	*	*	*	*	*	*
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	*	*	*	*	1133	1107	1124	29	*	*	*	*
LAB	*	*	*	*	1360	1077	1233	293	1090	1085	1456	251	1157	1127	1343	164
API	1107	1135	1509	231	1065	1129	1376	122	1116	1086	1299	162	1144	1127	1079	21
LABV	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	1531	1111	1068	251	1573	1149	1085	240	*	*	*	*
VEL	1391	1075	1145	257	1483	1124	1050	190	1382	1121	1003	95	*	*	*	*
ALV	1388	1137	1061	117	1358	1158	1239	187	*	*	*	*	*	*	*	*
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	1199	1175	1132	13	*	*	*	*	*	*	*	*
LAB	1015	1044	1146	49	1048	1074	1002	65	*	*	*	*	*	*	*	*
API	1190	1156	1227	70	*	*	*	*	1180	1252	1264	40	*	*	*	*
LABV	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Table A 1.g: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ʌ/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	1592	1207	1103	187	1699	1318	1447	340	*	*	*	*
VEL	*	*	*	*	*	*	*	*	1516	1221	1215	193	*	*	*	*
ALV	*	*	*	*	1406	1239	1163	61	442	1289	1326	127	1428	1255	1236	103
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	*	*	*	*	1152	1149	1283	91	*	*	*	*
LAB	*	*	*	*	1121	1181	1017	149	1038	1156	1331	38	1038	1105	1128	29
API	1141	1181	1439	145	1203	1209	1297	55	1184	1179	1261	58	1142	1174	1169	25
LABV									863	1068	1230	29				

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
VEL	1648	1319	1002	8	1404	1167	970	27	1469	1240	1110	66	*	*	*	*
ALV	1463	1279	1225	87	1404	1229	1020	23	1378	1265	1079	49	*	*	*	*
DEN	*	*	*	*	*	*	*	*	1272	1272	1234	25	*	*	*	*
LABD	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LAB	*	*	*	*	*	*	*	*	1164	1084	970	23	*	*	*	*
API	1188	1166	1103	27	1139	1107	986	59	*	*	*	*	*	*	*	*
LABV					1185	1123	968	62	909	1039	1172	2	*	*	*	*

Table A 1.h: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɒ/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	1436	1002	1058	327	1535	987	1218	519	*	*	*	*
VEL	*	*	*	*	*	*	*	*	1258	1071	1173	193	*	*	*	*
ALV	*	*	*	*	1361	1004	1016	246	1365	1081	1105	205	*	*	*	*
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	*	*	*	*	994	907	1197	251	*	*	*	*
LAB	*	*	*	*	1060	981	999	65	1004	940	1178	201	*	*	*	*
API	1029	1013	1333	224	1056	977	956	39	1040	1014	1308	213	1055	989	1132	139
LABV	812	948	1200	77	*	*	*	*	867	1016	1280	77	*	*	*	*

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	1433	1041	909	173	*	*	*	*	*	*	*	*
VEL	1252	997	1008	177	1197	1011	829	3	1206	1004	990	125	*	*	*	*
ALV	1368	1002	1072	291	1304	1053	899	65	1371	1071	938	111	*	*	*	*
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	*	*	*	*	955	958	986	17	*	*	*	*
LAB	841	915	1109	80	992	960	860	45	953	878	923	80	*	*	*	*
API	1007	965	1310	258	1061	970	862	11	1059	880	841	93	*	*	*	*
LABV	*	*	*	*	*	*	*	*	856	928	953	31	*	*	*	*

Table A 1.i: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɔ/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	*	*	*	*	1476	845	1434	813	*	*	*	*
VEL	*	*	*	*	*	*	*	*	1010	862	1218	336	*	*	*	*
ALV	*	*	*	*	1336	915	992	332	1260	834	1346	625	1280	812	1187	562
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	998	822	1513	578	*	*	*	*	884	786	1366	452	896	824	1059	205
LAB	*	*	*	*	831	831	1008	118	850	788	1192	429	*	*	*	*
API	*	*	*	*	*	*	*	*	1045	827	1204	397	949	987	1124	66
LABV	*	*	*	*	710	764	907	59	699	781	1412	366	*	*	*	*

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	*	*	*	*	1286	1067	1174	217	*	*	*	*
VEL	*	*	*	*	*	*	*	*	972	844	869	102	*	*	*	*
ALV	*	*	*	*	1397	908	893	316	1278	870	880	279	1295	943	831	160
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	963	802	909	179	944	893	963	81	838	746	797	95
LAB	753	777	1065	291	*	*	*	*	838	789	818	52	*	*	*	*
API	*	*	*	*	983	838	937	163	1018	889	813	35	115	935	772	11
LABV	*	*	*	*	663	764	851	9	767	820	809	43	*	*	*	*

Table A 1.j: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ʊ/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	1865	1893	1925	3	1743	1561	1322	38	1960	1841	1786	43	*	*	*	*
VEL	*	*	*	*	1236	1110	1052	45	1546	1472	1539	94	*	*	*	*
ALV	*	*	*	*	1672	1644	1596	13	1748	1746	1668	51	1584	1585	1487	66
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	*	*	*	*	1194	1202	1516	204	*	*	*	*
LAB	1055	1139	1569	231	1101	1013	999	49	1096	1323	1501	33	*	*	*	*
API	*	*	*	*	1278	1286	1281	9	1506	1577	1590	39	*	*	*	*
LABV	*	*	*	*	*	*	*	*	1107	1385	1610	35	*	*	*	*

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	1986	1858	1444	191	1593	1402	1278	45	2063	1992	1570	234
VEL	*	*	*	*	1695	1589	1113	247	1559	1259	1129	113	*	*	*	*
ALV	*	*	*	*	1754	1677	1379	147	1733	1537	1361	13	*	*	*	*
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	*	*	*	*	1194	949	954	167	*	*	*	*
LAB	1376	1565	1437	211	*	*	*	*	*	*	*	*	*	*	*	*
API	*	*	*	*	1414	1390	1167	133	*	*	*	*	*	*	*	*
LABV	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Table A 1.k: F2 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /u/.

	PAL				*VEL*				*ALV*				*DEN*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	2000	1795	1840	167	1907	1687	1445	15	1984	1824	1775	74	2017	1908	1794	3
VEL	*	*	*	*	*	*	*	*	1664	1534	1572	112	*	*	*	*
ALV	*	*	*	*	1763	1705	1593	36	1665	1711	1673	56	1710	1656	1576	17
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	*	*	*	*	1506	1567	1754	84	*	*	*	*
LAB	*	*	*	*	*	*	*	*	1372	1526	1658	15	*	*	*	*
API	*	*	*	*	1501	1462	1457	23	1492	1542	1695	69	1476	1496	1509	5
LABV	*	*	*	*	*	*	*	*	1064	1420	1547	153	*	*	*	*

	LABD				*LAB*				*API*				*LABV*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
PAL	*	*	*	*	2008	1827	1379	178	1702	1335	1215	165	2030	1811	1443	99
VEL	*	*	*	*	*	*	*	*	1455	1164	1033	107	*	*	*	*
ALV	*	*	*	*	1727	1667	1271	224	*	*	*	*	*	*	*	*
DEN	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LABD	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LAB	1382	1626	1505	243	*	*	*	*	*	*	*	*	*	*	*	*
API	1410	1506	1407	130	1443	1484	1276	166	*	*	*	*	*	*	*	*
LABV	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Table A 2.a: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /i/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	350	354	375	11	*	*	*	*	392	341	340	33	332	435	429	73	372	339	367	41
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Vel.stop	*	*	*	*	*	*	*	*	362	317	324	35	453	453	435	12	356	312	281	9	358	329	349	33	427	350	398	83
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	337	311	277	5	*	*	*	*	*	*	*	*
Alv.stop	367	352	357	13	*	*	*	*	405	341	463	124	298	297	288	5	391	326	319	39	398	425	459	5	408	348	358	47
Alv.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	401	377	326	18	457	447	321	77	441	392	365	15
Alv.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	353	343	327	4	379	417	414	27	385	356	357	20
Den	*	*	*	*	410	348	329	29	*	*	*	*	402	518	428	137	354	326	350	35	*	*	*	*	382	345	348	27
Labden	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	484	320	290	89	417	392	381	9	388	357	341	10
Lab.stop	372	352	346	9	*	*	*	*	380	349	373	37	*	*	*	*	377	340	304	1	392	351	365	37	375	337	347	32
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	387	373	354	3	372	391	373	25	410	340	355	57
Api	390	357	437	75	469	402	353	12	405	363	387	44	*	*	*	*	398	356	339	17	419	396	383	7	398	358	379	41
Labvel	*	*	*	*	*	*	*	*	386	360	322	8	363	347	339	5	385	336	308	14	405	395	333	35	*	*	*	*

Table A 2.a: F1 mean onset, midpoint and offset values and measures of deviation from linearity for /i/ (continued)

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	359	386	415	1	391	341	281	7	396	392	499	74	378	376	383	6	343	359	325	33
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Vel.stop	*	*	*	*	*	*	*	*	364	315	315	33	358	354	358	5	330	325	332	8	*	*	*	*
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	374	332	445	103	362	348	356	15	396	330	297	22	371	357	385	28	399	363	392	43	380	353	381	37
Alv.nas	*	*	*	*	436	392	394	31	492	396	315	10	479	401	403	53	413	404	388	5	451	428	384	14
Alv.fric	362	321	458	119	327	331	354	13	349	321	315	15	375	399	391	21	365	341	374	38	328	304	315	23
Den	*	*	*	*	379	336	348	37	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Labden	*	*	*	*	361	311	354	62	387	325	320	38	*	*	*	*	397	387	409	21	416	354	357	43
Lab.stop	398	387	378	1	372	344	460	96	377	350	341	12	*	*	*	*	447	369	387	64	404	358	365	35
Lab.nas	*	*	*	*	*	*	*	*	393	374	325	20	*	*	*	*	*	*	*	*	322	335	301	31
Api	435	384	372	26	408	359	394	56	384	337	330	27	393	380	385	12	412	377	377	23	418	375	356	16
Labvel	*	*	*	*	*	*	*	*	385	355	352	18	*	*	*	*	384	383	408	17	381	383	381	3

Table A 2.b: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɪ/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	399	386	372	1	431	489	447	67	395	370	338	5	491	465	416	15	414	398	386	3
Pal.gli	*	*	*	*	*	*	*	*	340	328	326	7	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Vel.stop	*	*	*	*	*	*	*	*	525	364	335	88	435	400	346	13	471	431	414	15	467	420	394	14	401	393	384	1
Vel.nas	*	*	*	*	*	*	*	*	289	288	285	1	*	*	*	*	*	*	*	*	457	495	463	47	402	436	398	48
Alv.stop	456	426	466	47	*	*	*	*	466	429	389	2	458	416	376	1	437	410	382	1	438	428	398	13	416	398	407	18
Alv.nas	511	493	475	0	*	*	*	*	*	*	*	*	*	*	*	*	486	480	407	45	455	516	429	99	445	434	395	19
Alv.fric	400	409	415	2	*	*	*	*	*	*	*	*	*	*	*	*	415	406	393	3	429	422	382	22	414	408	414	8
Den	407	418	509	57	*	*	*	*	428	406	360	16	482	519	483	49	*	*	*	*	458	469	445	23	410	419	414	9
Labden	444	420	472	51	*	*	*	*	462	443	370	36	459	440	381	27	460	458	446	7	505	480	442	7	443	433	438	10
Lab.stop	434	440	461	10	*	*	*	*	434	398	378	11	477	415	379	17	441	438	400	23	520	485	548	65	417	404	398	5
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	526	542	467	61	439	471	411	61	423	478	423	73	400	445	458	21
Api	473	485	547	33	*	*	*	*	485	459	405	19	486	481	405	47	493	475	440	11	494	490	431	37	463	453	458	10
Labvel	407	410	411	1	*	*	*	*	436	416	356	27	452	433	419	3	419	422	379	31	398	416	387	31	*	*	*	*

Table A 2.b: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɪ/ (continued).

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	384	435	448	25	472	452	432	0	386	432	461	11	410	445	468	8	410	379	382	23
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Vel.stop	*	*	*	*	393	414	444	6	441	440	452	9	*	*	*	*	418	449	492	8	*	*	*	*
Vel.nas	*	*	*	*	361	438	554	26	*	*	*	*	*	*	*	*	506	583	563	65	*	*	*	*
Alv.stop	408	390	430	39	463	440	452	23	415	431	403	29	464	447	416	9	461	462	476	9	*	*	*	*
Alv.nas	*	*	*	*	459	541	450	115	441	404	383	11	502	545	496	61	425	498	492	53	*	*	*	*
Alv.fric	*	*	*	*	414	404	420	17	499	440	410	19	437	496	502	35	440	436	444	8	*	*	*	*
Den	*	*	*	*	*	*	*	*	*	*	*	*	483	510	492	30	507	508	469	27	*	*	*	*
Labden	*	*	*	*	425	432	468	19	437	470	404	66	*	*	*	*	496	464	487	37	*	*	*	*
Lab.stop	*	*	*	*	501	499	660	109	417	415	374	26	*	*	*	*	454	472	476	9	*	*	*	*
Lab.nas	522	504	505	13	*	*	*	*	*	*	*	*	*	*	*	*	453	509	475	60	*	*	*	*
Api	420	442	445	13	435	446	455	1	444	451	404	36	471	502	453	53	450	486	483	26	401	403	385	13
Labvel	413	424	444	6	*	*	*	*	469	447	395	20	415	445	461	9	437	448	444	10	486	403	394	49

Table A 2.c: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity for (Dev) /ɛ/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	472	587	415	191	*	*	*	*	446	634	458	243	414	577	507	155	471	583	672	15
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	462	506	508	28	*	*	*	*	404	633	706	104
Vel.stop	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	459	540	486	90	488	590	479	142	424	585	591	103
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	497	690	454	286	*	*	*	*
Alv.stop	*	*	*	*	*	*	*	*	577	586	463	88	*	*	*	*	534	590	555	61	612	659	543	109	561	627	619	49
Alv.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	428	618	371	291	464	566	527	94	559	626	676	11
Alv.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	486	619	467	190	493	585	413	176	521	586	634	11
Den	*	*	*	*	*	*	*	*	624	623	442	120	*	*	*	*	*	*	*	*	564	632	514	124	*	*	*	*
Labden	*	*	*	*	*	*	*	*	604	633	481	121	*	*	*	*	539	576	458	103	517	627	552	123	528	621	621	62
Lab.stop	454	601	593	103	*	*	*	*	557	572	428	106	751	556	405	29	585	582	464	77	597	639	529	101	544	609	652	15
Lab.nas	550	566	417	110	*	*	*	*	*	*	*	*	*	*	*	*	578	594	469	94	545	629	512	134	*	*	*	*
Api	572	577	608	17	*	*	*	*	588	597	492	76	533	660	540	143	561	572	458	83	586	649	545	111	570	602	613	14
Labvel	494	579	620	29	*	*	*	*	*	*	*	*	*	*	*	*	554	621	524	109	542	565	485	69	*	*	*	*

Table A 2.c: Mean F1 onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɛ/ (continued).

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	429	679	695	156	*	*	*	*	*	*	*	*	490	586	616	44	*	*	*	*
Pal.gli	*	*	*	*	499	640	778	2	*	*	*	*	*	*	*	*	479	577	599	51	*	*	*	*
Vel.stop	468	631	554	160	*	*	*	*	576	608	519	81	702	642	622	27	537	696	699	104	*	*	*	*
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	*	*	*	*	490	601	585	85	509	661	576	158	582	655	615	75	640	642	619	17	*	*	*	*
Alv.nas	*	*	*	*	458	612	525	161	*	*	*	*	540	619	470	152	572	743	706	139	*	*	*	*
Alv.fric	*	*	*	*	486	584	504	119	*	*	*	*	497	612	481	164	495	633	613	105	*	*	*	*
Den	*	*	*	*	*	*	*	*	*	*	*	*	550	662	643	87	554	604	606	32	*	*	*	*
Labden	562	616	516	103	518	568	520	65	*	*	*	*	*	*	*	*	519	595	585	57	*	*	*	*
Lab.stop	583	675	653	76	547	595	469	116	*	*	*	*	*	*	*	*	512	586	589	47	*	*	*	*
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	533	729	673	168	533	663	595	132	*	*	*	*
Api	547	592	505	88	598	671	650	63	567	636	561	96	520	628	614	81	*	*	*	*	*	*	*	*
Labvel	479	550	477	96	*	*	*	*	535	583	470	107	*	*	*	*	512	609	620	57	*	*	*	*

Table A 2.d: F1 mean onset (On), midpoint (Mid) and offset values (Off) and measures of deviation from linearity (Dev) for /a/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	432	768	578	351	*	*	*	*	545	732	603	211	417	738	725	223	*	*	*	*
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	602	697	588	136
Vel.stop	642	759	680	131	*	*	*	*	664	630	424	115	742	727	376	224	730	765	592	139	661	769	615	175	569	563	411	77
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	531	662	471	215	*	*	*	*
Alv.stop	*	*	*	*	*	*	*	*	692	761	442	259	*	*	*	*	589	756	610	209	581	780	639	227	*	*	*	*
Alv.nas	471	712	683	180	*	*	*	*	*	*	*	*	*	*	*	*	592	752	618	196	537	690	516	218	*	*	*	*
Alv.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Den	*	*	*	*	*	*	*	*	*	*	*	*	668	734	569	154	574	713	665	125	*	*	*	*	*	*	*	*
Labden	638	798	830	85	*	*	*	*	543	775	558	299	588	767	605	227	644	771	602	197	*	*	*	*	796	790	780	3
Lab.stop	*	*	*	*	*	*	*	*	616	790	542	281	624	805	716	180	671	751	541	193	646	771	526	247	*	*	*	*
Lab.nas	566	786	711	197	*	*	*	*	*	*	*	*	731	830	767	108	601	769	493	296	591	739	576	207	576	725	619	170
Api	573	763	734	146	*	*	*	*	675	784	616	185	683	783	579	203	622	741	597	175	712	771	592	159	658	772	648	159
Labvel	*	*	*	*	*	*	*	*	*	*	*	*	679	765	704	98	*	*	*	*	*	*	*	*	*	*	*	*

Table A 2.d: Mean F1 onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /a/ (continued).

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	598	761	645	186	589	751	642	181	587	695	683	80	*	*	*	*
Pal.gli	*	*	*	*	*	*	*	*	576	797	727	194	*	*	*	*	*	*	*	*	*	*	*	*
Vel.stop	611	748	697	125	*	*	*	*	678	787	612	189	731	781	646	123	588	715	705	91	*	*	*	*
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	*	*	*	*	518	704	678	141	667	781	545	233	530	728	663	175	567	848	761	245	*	*	*	*
Alv.nas	540	724	736	115	*	*	*	*	544	789	666	245	699	756	696	78	371	879	764	415	*	*	*	*
Alv.fric	*	*	*	*	*	*	*	*	564	770	583	262	554	723	648	163	*	*	*	*	*	*	*	*
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Labden	*	*	*	*	*	*	*	*	*	*	*	*	637	800	635	219	546	741	645	194	*	*	*	*
Lab.stop	775	785	733	41	*	*	*	*	*	*	*	*	787	793	615	123	690	747	697	71	*	*	*	*
Lab.nas	662	790	768	100	*	*	*	*	*	*	*	*	506	750	662	221	663	754	741	69	*	*	*	*
Api	*	*	*	*	726	790	748	71	629	767	603	201	682	787	612	187	661	740	703	77	*	*	*	*
Labvel	*	*	*	*	*	*	*	*	*	*	*	*	624	762	663	158	*	*	*	*	*	*	*	*

Table A 2.e: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɜ/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	368	508	452	131	*	*	*	*	436	580	454	180	399	564	509	147	*	*	*	*
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	356	568	594	124	*	*	*	*
Vel.stop	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	465	554	483	107	*	*	*	*	607	563	411	72
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	469	589	440	179	*	*	*	*	*	*	*	*
Alv.stop	*	*	*	*	*	*	*	*	636	556	483	5	*	*	*	*	447	510	457	77	537	591	625	13	*	*	*	*
Alv.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	516	587	644	9
Alv.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	453	563	486	11	*	*	*	*	478	449	424	3
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	459	507	479	51	*	*	*	*	502	532	415	98
Labden	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	452	541	395	157	451	555	583	51	497	554	536	50
Lab.stop	*	*	*	*	*	*	*	*	466	529	380	141	*	*	*	*	536	552	520	32	514	533	513	26	602	459	599	69
Lab.nas	430	559	366	215	*	*	*	*	*	*	*	*	*	*	*	*	457	528	413	124	*	*	*	*	539	576	603	7
Api	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	502	552	385	145	495	581	477	127	*	*	*	*
Labvel	514	554	633	26	*	*	*	*	449	548	442	137	*	*	*	*	459	542	411	143	*	*	*	*	*	*	*	*

Table A 2.e: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɜ/ . (continued)

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	429	579	406	215	*	*	*	*	*	*	*	*	*	*	*	*
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Vel.stop	*	*	*	*	*	*	*	*	516	580	462	121	*	*	*	*	455	588	632	59	*	*	*	*
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	*	*	*	*	*	*	*	*	543	555	489	52	599	572	563	12	*	*	*	*	*	*	*	*
Alv.nas	506	586	713	31	466	491	402	76	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.fric	*	*	*	*	484	552	514	71	457	536	517	65	494	572	485	110	420	517	443	114	459	549	605	23
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Labden	489	548	494	75	*	*	*	*	503	549	428	111	499	580	509	101	507	548	543	31	*	*	*	*
Lab.stop	495	547	622	15	541	524	529	15	562	557	496	37	578	566	534	13	681	620	614	37	*	*	*	*
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Api	479	535	529	41	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Labvel	463	554	515	87	433	534	436	132	*	*	*	*	455	547	537	68	437	502	547	13	*	*	*	*

Table A 2.f: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɑ/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	430	621	672	93	*	*	*	*	586	658	457	182	536	761	642	229	*	*	*	*
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	616	713	549	174	*	*	*	*	*	*	*	*
Vel.stop	*	*	*	*	*	*	*	*	644	691	512	151	*	*	*	*	564	681	606	128	683	728	535	159	606	704	771	21
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	715	771	697	87	*	*	*	*	*	*	*	*
Alv.stop	547	713	651	152	*	*	*	*	533	715	559	225	*	*	*	*	557	650	636	71	495	782	516	369	606	724	611	154
Alv.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	607	733	712	98
Alv.fric	508	750	459	355	*	*	*	*	*	*	*	*	*	*	*	*	496	726	539	278	519	657	570	150	*	*	*	*
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	564	677	394	264	*	*	*	*	*	*	*	*
Labden	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	581	736	501	260	573	736	556	229	*	*	*	*
Lab.stop	*	*	*	*	*	*	*	*	654	699	554	127	637	707	668	73	*	*	*	*	*	*	*	*	645	682	675	29
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	489	728	680	191	*	*	*	*	530	727	762	108	*	*	*	*
Api	636	721	502	203	*	*	*	*	615	693	448	215	*	*	*	*	581	677	499	183	659	731	555	165	652	706	734	17
Labvel	*	*	*	*	*	*	*	*	671	644	397	147	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Table A 2.f: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɑ/ (continued).

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	447	699	554	265	510	710	662	165	494	684	639	157	*	*	*	*
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Vel.stop	*	*	*	*	634	715	646	100	575	653	572	81	501	758	720	197	634	702	596	116	676	662	499	99
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	*	*	*	*	525	698	561	207	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.nas	*	*	*	*	421	734	737	207	*	*	*	*	679	745	678	89	*	*	*	*	*	*	*	*
Alv.fric	*	*	*	*	523	732	580	241	516	674	484	232	519	734	477	315	*	*	*	*	*	*	*	*
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Labden	*	*	*	*	*	*	*	*	*	*	*	*	607	678	527	148	604	750	793	69	*	*	*	*
Lab.stop	648	748	691	105	659	695	724	5	553	694	524	207	*	*	*	*	526	690	552	201	*	*	*	*
Lab.nas	*	*	*	*	498	671	486	239	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Api	597	653	535	116	594	733	749	82	*	*	*	*	620	696	594	119	577	704	700	87	*	*	*	*
Labvel	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Table A 2.g: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ʌ/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	*	*	*	*	501	652	456	231	436	639	386	304	612	689	699	45	461	574	693	4
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	607	695	488	197	*	*	*	*	*	*	*	*	*	*	*	*
Vel.stop	493	657	656	110	*	*	*	*	*	*	*	*	*	*	*	*	707	659	539	48	563	695	527	200	646	697	654	63
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	*	*	*	*	*	*	*	*	632	688	536	139	643	734	477	232	535	588	505	91	570	645	482	159	480	520	425	90
Alv.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	561	670	569	140	646	675	527	118	*	*	*	*
Alv.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	571	626	680	1	533	707	576	203	*	*	*	*
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	505	676	556	194	*	*	*	*
Labden	*	*	*	*	*	*	*	*	*	*	*	*	559	643	341	257	626	675	502	148	619	701	600	122	*	*	*	*
Lab.stop	*	*	*	*	*	*	*	*	*	*	*	*	592	717	411	287	547	604	630	21	**558**	695	491	227	545	682	739	53
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	525	713	428	315	527	658	571	145	623	710	589	139	511	699	705	121
Api	594	691	698	60	*	*	*	*	606	658	469	161	607	651	609	57	*	*	*	*	650	710	573	131	617	667	639	52
Labvel	*	*	*	*	*	*	*	*	*	*	*	*	582	588	429	110	*	*	*	*	536	650	563	134	*	*	*	*

Table A 2.g: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /^/ (continued).

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	490	673	597	173	604	679	466	192	*	*	*	*	*	*	*	*	*	*	*	*
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Vel.stop	653	673	847	103	590	613	536	67	663	691	633	57	678	663	482	111	647	644	606	23	*	*	*	*
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	*	*	*	*					571	702	608	150	642	739	586	167	528	617	544	108	*	*	*	*
Alv.nas	541	718	699	131	559	695	713	79	444	721	393	403	523	692	489	248	567	680	677	77	*	*	*	*
Alv.fric	*	*	*	*	581	688	740	37	550	621	483	139	533	611	491	132	460	607	696	39	*	*	*	*
Den	515	686	506	87	*	*	*	*	*	*	*	*	*	*	*	*	561	608	599	37	*	*	*	*
Labden	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	582	624	514	101	*	*	*	*
Lab.stop	*	*	*	*	*	*	*	*	671	632	506	58	601	662	550	115	597	628	593	44	*	*	*	*
Lab.nas	580	675	587	122	451	680	601	205	687	648	502	71	562	692	586	157	445	650	637	145	*	*	*	*
Api	601	638	573	68	578	631	540	96	644	651	479	119	665	668	522	99	650	681	693	13	*	*	*	*
Labvel	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	541	635	649	53	*	*	*	*

Table A 2.h: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɒ/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	544	623	674	19	*	*	*	*					467	633	558	161	*	*	*	*
Pal.gli	*	*	*	*	*	*	*	*	585	640	354	227	*	*	*	*	549	666	511	181	636	683	497	155	*	*	*	*
Vel.stop	*	*	*	*	*	*	*	*	539	592	433	141	*	*	*	*	516	581	549	65	608	594	463	78	551	523	604	73
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	381	487	373	147	*	*	*	*
Alv.stop	*	*	*	*	*	*	*	*	520	596	461	141	*	*	*	*	596	599	496	71	562	600	510	48	*	*	*	*
Alv.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	529	630	520	141	484	526	467	67	*	*	*	*
Alv.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	459	614	658	74	493	577	487	116	482	633	683	67
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Labden	*	*	*	*	*	*	*	*	560	619	421	171	*	*	*	*	502	523	523	14	550	584	440	119	*	*	*	*
Lab.stop	*	*	*	*	*	*	*	*	567	626	582	69	592	670	549	133	496	557	605	9	679	671	511	101	557	593	621	5
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	555	679	539	176	430	522	510	69	*	*	*	*
Api	478	591	517	125	*	*	*	*	566	593	486	89	563	635	434	182	540	590	480	107	566	612	518	93	563	638	711	1
Labvel	506	629	733	13	*	*	*	*	*	*	*	*	*	*	*	*	508	561	532	55	472	584	522	116	*	*	*	*

Table A 2.h: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɒ/ (continued).

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Pal.gli	*	*	*	*	478	499	581	41	536	621	517	126	*	*	*	*	*	*	*	*	*	*	*	*
Vel.stop	*	*	*	*	615	618	679	39	618	591	444	80	656	643	494	91	581	591	539	41	*	*	*	*
Vel.nas	*	*	*	*	434	678	515	271	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	*	*	*	*	567	629	614	51	653	617	514	45	604	638	468	136	478	597	550	111	*	*	*	*
Alv.nas	*	*	*	*	573	640	521	124	453	650	432	277	601	431	296	23	585	654	617	71	*	*	*	*
Alv.fric	*	*	*	*	489	605	677	29	*	*	*	*	*	*	*	*	510	602	541	102	*	*	*	*
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	569	565	572	7	*	*	*	*
Labden	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	533	565	580	11	*	*	*	*
Lab.stop	*	*	*	*	*	*	*	*	595	603	465	97	*	*	*	*	569	553	527	7	*	*	*	*
Lab.nas	*	*	*	*	400	488	614	25	*	*	*	*	557	648	607	88	477	561	536	73	*	*	*	*
Api	584	620	744	59	579	605	733	68	567	590	460	102	629	668	630	51	507	488	476	5	*	*	*	*
Labvel	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	512	559	547	39	*	*	*	*

Table A 2.i: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɔ/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	460	465	616	97	523	443	716	235	407	448	503	9
Pal.gli	*	*	*	*	*	*	*	*	406	498	550	27	*	*	*	*	448	477	471	23	*	*	*	*	500	531	598	24
Vel.stop	472	474	668	128	*	*	*	*	*	*	*	*	*	*	*	*	453	471	454	24	478	510	524	12	482	459	462	17
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	412	459	413	62	*	*	*	*	*	*	*	*
Alv.stop	*	*	*	*	*	*	*	*	447	494	420	81	*	*	*	*	519	463	527	80	431	516	534	45	443	471	397	68
Alv.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	418	458	445	35	*	*	*	*	*	*	*	*
Alv.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	431	456	652	114	*	*	*	*	465	453	684	162
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	444	461	445	22	*	*	*	*	*	*	*	*
Labden	*	*	*	*	524	447	462	61	*	*	*	*	*	*	*	*	426	431	401	23	418	465	469	29	441	469	612	77
Lab.stop	470	508	652	71	*	*	*	*	453	482	751	160	678	632	371	143	448	464	483	2	418	434	570	80	436	416	429	22
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	391	482	391	121	348	462	503	49	446	538	548	55
Api	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	449	462	471	3	490	505	452	45	478	473	448	13
Labvel	*	*	*	*	*	*	*	*	404	444	388	64	*	*	*	*	399	452	563	39	*	*	*	*	*	*	*	*

Table A 2.i: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ɔ/ (continued).

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	412	471	498	21	*	*	*	*
Vel.stop	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	456	457	461	2	425	457	404	57
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	453	464	510	23	*	*	*	*	422	448	382	61	*	*	*	*	458	452	463	11	430	458	391	63
Alv.nas	546	520	641	98	*	*	*	*	*	*	*	*	445	550	429	151	462	468	476	1	470	534	500	65
Alv.fric	*	*	*	*	*	*	*	*	434	451	355	75	*	*	*	*	438	441	443	1	*	*	*	*
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	459	456	408	30	*	*	*	*
Labden	442	461	473	5	*	*	*	*	382	452	433	59	510	470	437	5	482	484	490	3	435	425	409	4
Lab.stop	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	437	453	425	29	*	*	*	*
Lab.nas	469	486	404	66	400	488	614	92	*	*	*	*	397	488	356	149	408	462	448	45	*	*	*	*
Api	478	486	416	52	424	444	327	91	460	445	389	27	487	529	294	185	504	481	450	5	476	508	399	94
Labvel	*	*	*	*	427	488	590	27	391	472	516	25	383	466	440	73	413	458	452	34	*	*	*	*

Table A 2.j: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ʊ/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	408	466	407	78	*	*	*	*	500	400	376	51	*	*	*	*	*	*	*	*
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	369	371	367	4	354	344	341	5	367	392	409	5
Vel.stop	*	*	*	*	*	*	*	*	480	441	361	27	*	*	*	*	440	431	408	9	*	*	*	*	351	352	337	11
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	*	*	*	*	*	*	*	*	471	417	398	23	*	*	*	*	413	418	376	31	*	*	*	*	*	*	*	*
Alv.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Labden	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	428	436	404	27	*	*	*	*	*	*	*	*
Lab.stop	436	446	440	11	*	*	*	*	438	463	408	53	*	*	*	*	456	443	380	33	*	*	*	*	425	450	536	41
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Api	*	*	*	*	*	*	*	*	444	449	417	25	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Labvel	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	392	387	361	14	*	*	*	*	*	*	*	*

Table A 2.j: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /ʊ/ (continued).

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	*	*	*	*	551	430	518	139	386	407	447	13	*	*	*	*
Pal.gli	*	*	*	*	*	*	*	*	*	*	*	*	342	341	330	7	414	444	468	4	*	*	*	*
Vel.stop	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	509	497	469	11	*	*	*	*
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	491	423	439	56	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	388	405	431	6	*	*	*	*
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Labden	*	*	*	*	*	*	*	*	459	441	406	11	*	*	*	*	539	434	419	60	*	*	*	*
Lab.stop	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	429	440	453	1	*	*	*	*
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Api	*	*	*	*	*	*	*	*	*	*	*	*	430	463	507	7	*	*	*	*	*	*	*	*
Labvel	*	*	*	*	*	*	*	*	*	*	*	*	377	437	412	57	402	417	428	3	*	*	*	*

Table A 2.k: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /u/.

	Pal.fric				*Pal.gli*				*Vel.stop*				*Vel.nas*				*Alv.stop*				*Alv.nas*				*Alv.fric*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	367	357	363	11	382	421	529	46	382	353	370	31
Pal.gli	*	*	*	*	*	*	*	*	326	353	340	27	*	*	*	*	348	346	353	6	*	*	*	*	342	343	351	5
Vel.stop	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	391	426	415	31	356	361	362	3	421	367	346	22
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	*	*	*	*	332	337	328	9	364	361	342	11	*	*	*	*	*	*	*	*	449	357	397	88	388	338	379	61
Alv.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	392	488	457	85	*	*	*	*
Alv.fric	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	394	368	365	15	387	390	364	19	377	368	368	6
Den	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Labden	351	322	379	57	*	*	*	*	*	*	*	*	*	*	*	*	398	375	336	11	*	*	*	*	389	353	350	22
Lab.stop	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	395	341	311		379	348	335	12	*	*	*	*
Lab.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	373	359	276	46	365	381	354	29	*	*	*	*
Api	363	351	350	7	*	*	*	*	400	356	484	115	*	*	*	*	372	354	332	3	*	*	*	*	380	346	356	29
Labvel	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	381	371	323	25	*	*	*	*

Table A 2.k: F1 mean onset (On), midpoint (Mid) and offset (Off) values and measures of deviation from linearity (Dev) for /u/ (continued).

	Den				*Labden*				*Lab.stop*				*Lab.nas*				*Api*				*Labvel*			
	On	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*	*On*	*Mid*	*Off*	*Dev*
Pal.fric	*	*	*	*	*	*	*	*	493	342	345	103	*	*	*	*	354	385	428	8	*	*	*	*
Pal.gli	340	356	445	49	*	*	*	*	367	342	327	7	357	406	405	33	390	410	412	12	376	370	365	1
Vel.stop	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	389	387	403	12	*	*	*	*
Vel.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.stop	422	359	372	51	420	338	334	52	*	*	*	*	358	380	367	23	356	648	391	34	362	347	342	7
Alv.nas	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Alv.fric	373	329	350	43	*	*	*	*	349	349	354	3	414	384	307	31	*	*	*	*	*	*	*	*
Den	515	373	362	87	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Labden	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Lab.stop	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Lab.nas	*	*	*	*	416	408	358	28	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Api	376	368	359	1	395	368	400	39	354	344	327	5	406	407	417	6	*	*	*	*	430	369	326	12
Labvel	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	394	443	475	11	*	*	*	*

References

Ainsworth, W. (1975). *Intrinsic and extrinsic factors in vowel judgements.* In: Auditory Analysis and Perception of Speech, Fant, G. and Tatham, M. (Eds.). Academic, London, 103-113.

Alfonso, P. J. and Baer, T. (1982). Dynamics of vowel articulation. *Language and Speech,* 25, 151-173.

Altman, G. and Cutler, D. M. (1989). Lexical stress and lexical discriminability: Stressed syllables are more informative, but why? *Computer Speech and Language*, 3, 265-275.

Anderson S. R. (1982). The analysis of French schwa: or how to get something for nothing. *Language,* 58, 535-573.

Archangeli, D. (1984). Underspecification in Yawelmani Phonology and Morphology. *PhD thesis*, MIT. Published 1988, New York: Garland.

Archangeli, D. (1988). Aspects of Underspecification Theory. *Phonology* 5. 183-205.

Assman, P. F., Nearey, T. M. and Hogan, J. T. (1982). Vowel Identification: orthographic, perceptual, and acoustic aspects. *Journal of the Acoustical Society of America,* 71 (4), 975-989.

Bagshaw, P. C., Hiller, S. M. and Jack, M. A. (1993). Enhanced pitch tracking and the processing of F0 contours for computer aided teaching. *Proceedings of the Third European Conference on Speech Communciation and Technology*, 2, Berlin, 103-106.

Barry, W. J. (1992). Comments on Chapter 2: "Targetless" schwa: an articulatory analysis. In: Docherty, G. and Ladd, R. (Eds.), *Gesture, Segment, Prosody. Papers in Laboratory Phonology II.* 65-67, Cambridge University Press.

Becker, R. A. and Chambers, J. M. (1984). S: An Interactive Environment for Data Analysis and Graphics. Belmont, California: Wadsworth.

Bell-Berti, F. and Harris, K. S. (1976). Some aspects of coarticulation. *Status Report on Speech Research,* SR-45/46, Haskins Laboratories.

Benguerel, A. P. and Cowan, H. A. (1974). Coarticulation of upper lip protrusion in French. *Phonetica,* 30, 41-55.

Bladon, R. A. W. (1982). Arguments against formants in the auditory representation of speech. In Carlson, R. and Granstrom, B. (Eds.), *The representation of speech in the peripheral auditory system*, 95-102. Amsterdam: Elsevier Biomedical Press.

Bladon, R. A. W., Henton, C. G. and Pickering, J. B. (1984). Towards an auditory theory of speaker normalisation. *Language Communication*, 4, 59-69.

Blumstein, S. E. and Stevens, K. N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America*, 66, 1001-1017.

BMDP Statistical Software Manual. BMDP Statistical Software, Los Angeles. University of California Press, Berkeley, 94720.

Bolinger, D. (1975). *Aspects of Language*, Harcourt Brace Jovanovich, New York.

Bolinger, D. (1985). Two views of accent. *Journal of Linguistics,* 21, 79-123.

Browman, C. P. and Goldstein, L. (1984). Dynamic modeling of phonetic structure. *Status Report on Speech Research*, SR-79/80, Haskins Laboratories.

Browman, C. P. and Goldstein, L. (1986) Towards an articulatory phonology. *Phonology Yearbook* 3, 219-252.

Browman, C. P. and Goldstein, L. (1989). Gestural structures and phonological patterns. *Status Report on Speech Research,* SR-97/80, Haskins Laboratories.

Browman, C. P. and Goldstein, L. (1992). Targetless schwa: an articulatory analysis. In Docherty, G. J. and Ladd, R. (Eds.), *Gesture, Segment, Prosody. Papers in Laboratory Phonology II*, 26-67. Cambridge University Press.

Byrd, D. (1994). Articulatory Timing in English Consonant Sequences. *PhD thesis*, UCLA.

Carney, P. J. and Moll, K. L. (1971). A cinefluorographic investigation of fricative consonant-vowel coarticulation. *Phonetica*, 23, 193-202.

Carter, D. M. (1987). An information-theoretic analysis of phonetic dictionary access. *Computer Speech and Language*, 2, 1-11.

Chiba, T. S. and Kajiyama, M. (1941). The vowel: Its Nature and Structure. *Tokyo-Kaiseikan,* Tokyo.

Chistovich, L. A. (1985). Central auditory processing of peripheral vowel spectra. *Journal of the Acoustical Society of America,* 77 (3), 789-805.

Chistovich, L. A. and Lublinskaya, V. V. (1979). The 'Center of gravity' effect in vowel spectra and critical distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli. *Journal of Hearing Research*, 1, 185-195.

Choi, J. D. (1992). Phonetic underspecification and target-interpolation: An acoustic study of Marshallese vowel allophony. *UCLA Working Papers in Phonetics*, 82, 1-133.

Chomsky, N. and Halle, M. (1968). Sound Pattern of English. *Harper & Row*; New York.

Cohen, J. and Cohen, P. (1983). Applied Multiple Regression/Correlational Analysis for the Behavioural Sciences. *Lawrence Erlbaum Associates*; Hillsdale, New Jersey, London.

Couper-Kuhlen, E. (1993). English Speech Rythmn: form and function in everyday verbal interactions. *John Benjamins*, Amsterdam.

Crowe, A. (1987). Globally Optimising Formant Tracker Using Generalised Centroids. Electronics Letters, Vol 23, No 19, 1019-1020.

Dalby, R. M. (1984). Phonetic Structure of Fast Speech Processes in American English. *PhD thesis*, Indiana University.

Daniloff, R. G. and Hammarberg, R. E. (1973). On Defining Coarticulation. *Journal of Phonetics* 1, 239-248.

Delattre, P., Liberman, A. M., Cooper, F. S. and Gerstman, L. J. (1952). An experimental study of the acoustic determinants of vowel colour: Observations on one and two-formant vowels synthesised from spectrographic patterns. *Word*, 8, 195-210.

Delattre, P. (1969). An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics*, 7, 295-325.

Den Os, E. (1988). Rhythm and tempo of Dutch and Italian. *Phd thesis*, University of Utrecht.

Di Benedetto, M. (1987). An acoustical and perceptual study on vowel height. *PhD thesis*, Universita degli Studi di Roma 'La Sapienza', Rome, Italy.

Di Benedetto, M. (1989). Vowel Representation: Some observations on temporal and spectral properties of the first formant frequency. *Journal of the Acoustical Society of America*, 86 (1), 55-66.

Disner, S. F. (1980). Evaluation of vowel normalisation procedures. *Journal of the Acoustical Society of America*, 67(1), 253-261.

Eimas, P. D. and Miller, J. L. (1978). Selective adaptation of speech and visual perception. In: Walk, R. D. and Pick, H. L. (Eds.), *Perception and Experience*. New York: Plenum Press.

Engstrand, O. (1988). Articulatory correlates of stress and speaking rate in Swedish VCV utterances. *Journal of the Acoustical Society of America*, 83 (4), 1863-1875.

Fant, G. (1956). On the predictability of formant levels and spectrum envelopes from formant frequencies. In Halle, M., Lunt, H. and Maclean, H. (Eds.). *For Roman Jackobson*. The Hague: Mouton.

Fant, G. (1960). *Acoustic Theory of Speech Production*. Mouton, The Hague.

Fant, G. (1973). *Speech, Sounds and Features*. Cambridge, Mass. MIT Press.

Fant, G. (1983). Feature analysis of Swedish vowels - a revisit, *Speech transmission laboratory:* Q. Prog. Stat. Rep. 2-3, 1-19.

Fant, G., Carlson, R., and Granstrom, B. (1974). The (e)-(ø) ambiguity, *Proc Speech Communication Semin.*, Stockholm 3, 117-121.

Farnetani, E. and Faber, A. (1992). Tongue-jaw coordination in vowel production: Isolated words versus connected speech. *Speech Communication*, 11, 401-410.

Farnetani, E., Padova, C. N. R. and Recasens, D. (1993). Anticipatory consonant-to-vowel coarticulation in the production of VCV sequences in Italian. *Language and Speech*, 36 (2,3), 279-302.

Flanagan, J. L. (1955). A difference limen for vowel formant frequency. *Journal of the Acoustical Society of America* 27, 306-10.

Flege, J. E. (1988). Effects of speaking rate on tongue position and velocity of movement in vowel production. *Journal of the Acoustical Society of America*, 84 (3), 901-916.

Fourakis, M. (1991). Tempo,stress and vowel reduction in American English. *Journal of the Acoustical Society of America*, 90 (4), 1816-1827.

Fowler, C. A. (1981). Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech and Hearing Research*, 24, 127-139.

Fowler, C. A. (1983). Converging sources of evidence for spoken and perceived rhythms of speech: cyclic production of vowels in sequences of monosyllabic stress feet. *Journal of Experimental Psychology*: General 112, 386-412.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14, 3-28.

Fry, D. B. (1958). Experiments in the Perception of Stress, *Language and Speech* 1, 126-152.

Fry, D. B. (1981). *The Physics of Speech*. Cambridge: Cambridge University Press.

Fudge, E. (1984). *English Word-Stress*. London: Allen and Unwin.

Gay, T. (1974). A cinefluorographic study of vowel production. *Journal of Phonetics*, 2, 255-266.

Gay, T. and Hirose, H. (1973). Effect of speaking rate on labial consonant production: a combined electromyographic high-speed motion picture study. *Phonetica*, 27, 203-213.

Gay, T., Ushijima, H., Hajime Hirose and Cooper, F. S. (1974). Effect of speaking rate on labial consonant-vowel articulation. *Journal of Phonetics*, 2, 47-63.

Gay, T. (1977). Articulatory movements in VCV sequences. *Journal of the Acoustical Society of America*, 62, 183-193.

Gay, T. (1978). Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Scoiety of America*, 63, 223-230.

Gerstman, L. H. (1968). Classification of self-normalised vowels. IEEE Trans. Audio Electroacoust. AU-16, 78-80.

Harmegnies, B. and Poch-Olive. (1992). A study of style-induced vowel variability: Laboratory versus spontaneous speech in Spanish. *Speech Communication* 11, 429-437, North Holland.

Harris, K. S. (1971). Vowel stress and articulatory reorganisation. *Haskins Laboratories Status Report on Speech Research*. SR-28, 167-178.

Harris, K. S. (1973). Stress and syllable duration change. *Haskins Laboratories Status Report on Speech Research*. SR-35/36, 31-38.

Harris, K. S. (1978). Vowel duration change and its underlying physiological mechanisms. *Language and Speech*, 21, 354-361.

Harris, K. S. (1984). Coarticulation as a component in articulatory description. *Status Report on Speech Research*, SR79/80, Haskins Laboratories.

Harris, K. S. and Umeda, N. (1974). Effect of speaking mode on temporal factors in speech. *Journal of the Acoustical Society of America*, 56 (3), 1016-1018.

Hayes, B. (1984). *A metrical theory of stress rules*. Garland Press, New York.

Henke, W. L. (1966). Dynamic Articulatory Model of Speech Production Using Computer Simulation. *PhD thesis*, MIT.

Henton, C. G. (1983). Changes in vowels of received pronunciation. *Journal of Phonetics*, 11, 353-371.

Hillenbrand, J. and Gayvert, R. T. Vowel classification based on fundamental and formant frequencies. *Journal of Speech and Hearing Research*, 36, 694-700.

House, A. S. (1961). On Vowel Duration in English, *Journal of the Acoustical Society of America*, 33, 1174-1178.

House, A. S. and Stevens, K. N. (1956). Analog Studies of the Nasalization of Vowels. *Journal of Speech and Hearing Disorders*, 21, No 2, 218-232.

Huang, C. B. (1985). Perceptual correlates of the tense/lax distinction in General American English. *Master's thesis*, MIT, Cambridge, MA, USA.

Huang, C. B. (1992). Modelling human vowel identification using aspects of formant trajectory and context. In: Tohkura, Y., Vatikiotis-Bateson, E. and Sagisaka, Y. (Eds.), *Speech Perception, Production and Linguistic Structure*, 43-61. IOS Press.

Huffman, M. K. (1986). Patterns of coarticulation in English. *Working Papers in Phonetics* 63, Universtity of California Linguistics Club.

James, M. (1985). *Classification Algorithms*. London: Collins.

Joos, M. (1948). Acoustic Phonetics. *Language*, 24, 1-136.

De Jong, K., Beckman, M. E. and Edwards, J. The Interplay between prosodic structure and coarticulation. *Language and Speech*, 36 (2,3), 197-212.

Kanter, C. and West, R. (1960). *Phonetics*. Harper.

Keating, P. A. (1988). Underspecification in phonetics. *Working papers in Phonetics* 69, University of California Linguistics Club.

Keating, P. A. (1990). The window model of coarticulation. In Kingston, J. and Beckman, M. E. (Eds.), *Papers in Laboratory Phonology I: Between Grammar and Physics of Speech*. Cambridge University Press, Cambridge.

Keating, P. Lindblom, B., Lubker, J. and Kreiman, J. (1990). Jaw position in English and Swedish VCV's. *UCLA Working Papers in Phonetics,* 74, 77-95.

Kelso, J. A., Saltzman, E. L. and Tuller, B. (1986). The dynamical perspective on speech production: data and theory. *Journal of Phonetics*, 14, 29-59.

Kent, R. D. and Moll, K. L. (1972). Tongue body articulation during vowel and diphthong gesture. *Folia Phoniatrica*, 25, 278-300.

Kent, R. D., Carney, P. J. and Severeid, L. R. (1974). Velar movement and timing: Evaluation of a model for binary control. *Journal of Speech and Hearing Research,* 17, 470-488.

Kerswill, P. E. (1985). A sociophonetic study of connected speech processes in Cambridge English: an outline and some results. *Cambridge Papers in Phonetics and Experimental Linguistics*, 9, University of Cambridge.

Kingston, J. (1992). Comments on Chapter 2: "Targetless" schwa: an articulatory analysis. In: Docherty, G. and Ladd, R. (Eds.), *Gesture, Segment, Prosody. Papers in Laboratory Phonology II*. 65-67, Cambridge University Press.

Kiparsky, P. (1985). Some consequences of lexical phonology. *Phonology* 2, 85-138.

Klatt, D. H. (1973). Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America*, 54 (4), 1102-1104.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics*, 3, 129-140.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59 (5), 1208-1221.

Koopmans-van Beinum, F. J. (1980). Vowel contrast reduction: an acoustic and perceptual study of Dutch vowels in various speech conditions, *Ph.D thesis*, University of Amsterdam,.

Koopmans-van Beinum, F. J. (1994). What's in a schwa? Durational and spectral analysis of natural continuous speech and diphones in Dutch. *Phonetica*, 51, 68-79.

Kondo, Y. (1995). Production of schwa by Japanese speakers of English: a crosslinguistic study of coarticulatory strategies. *PhD thesis*, University of Edinburgh.

Kozhevnikov, V. A. and Chistovich, L. A. (1965). *Artikulyatsiya i Vospriyatiye. Trans Speech: Articulation and Perception.* Vol. 30, 543.

Krull, D. (1987). Second-formant locus patterns as a measure of consonant-vowel coarticulation. *Phonetic Experimental Research, Institute of Linguistics, University of Stockholm (PERILUS)* V, 43-61.

Krull, D. (1989). Second formant locus patterns and consonant-vowel coarticulation in spontaneous speech. *Phonetic Experimental Research, Institute of Linguistics, University of Stockholm (PERILUS)* X, 87-108.

Kuehn, D. and Moll, K. (1972). Perceptual effects of forward coarticulation. *Journal of Speech and Hearing Research*, 15, 654-664.

Kuehn, D and Moll, K. (1976). A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics*, 4, 303-320.

Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press, Philadelphia.

Ladefoged, P. and Broadbent, D. (1957). Information conveyed by vowels. *Journal of Acoustical Society of America*, 29, 98-104.

Lamel, L. F., Kassel, R. H. and Seneff, S. (1986). Speech database development; design and analysis of the acoustic-phonetic corpus, pp. 100-109 of Baumann, L. S. (Ed.) *Proceedings of DARPA Speech Recognition Workshop*. Palo Alto, California, 19-20 February 1986. McLean, Virginia: Science Applications International Corporation.

Laver, J., Bennett, C., Cohen, I., Dalby, J., Davies, D., McAllister, M. (1988). *ATR/CSTR Speech Database Project Status Report No 1*, Centre for Speech technology research, University of Edinburgh.

Laver, J., Hieronymus, J., Bennett, C., Cohen, I., Dalby, J., Davies, D., Hiller, S., McAllister, M., Purves, E. M. (1989). *ATR/CSTR Speech Database Project Status report No 3*, Centre for Speech Technology Research, University of Edinburgh.

Lehiste, I. and Peterson, G. E. (1961). Transitions, glides and diphthongs. *Journal of the Acoustical Society of America*, 33, 268-277.

Lehiste, I. (1964). *Acoustical Characteristics of Selected English Consonants*. Mouton: The Hague.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Mass: MIT press.

Liberman, A. M. and Mattingly, I. G. (1985). The motor theory of speech revisited. *Cognition*, 21, 1-36.

Liberman, A. M. and Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8, 249-336.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America*, 32, 451-454.

Lieberman, P. (1984). *The Biology and Evolution of Language*. Harvard University Press, Cambridge, MA.

Lindau, M. and Ladefoged, P. (1990). Interarticulatory realationships in vowel production. *UCLA Working Papers.*

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 35, 1773-1781.

Lindblom, B. (1967). Vowel duration and a model of Lip Mandible Coordination. *Speech Transmission Lab. Quart. Progr. Status Rp.* 4/1967 (Royal Inst. Techn., Stockholm), 1-29.

Lindblom, B. and Studdert-Kennedy, M. (1967). On the role of formant transitions in vowel recognition. *Journal of the Acoustical Society of America*, 42 (4), 830-843.

Lindblom, B. and Sundberg, J. (1971). Acoustical consequences of lip, tongue, jaw and larynx movement. *Journal of the Acoustical Society of America*, 50, 116-1179.

Lindblom, B. (1983) Economy of Speech gestures. In: MacNeilage, P. F. (Ed.), *Production of Speech*. Springer; New York. 217-245.

Lindblom, B. and Moon, S.-J. (1988). Formant undershoot in clear and citation-form speech. *Phonetic Experimental Research, Institute of Linguistics, University of Stockholm (PERILUS), VIII*, 21-33.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the HH theory. In: Hardcastle, W. J. and Marchal, A. (Eds.), *Speech Production and Speech Modeling*, 403-439. Kluwer Academic Publishers.

Lindblom, B., Brownlee, S., Davis, B. and Moon, S. J. (1992). Speech transforms. *Speech Communication*, 11, 357-368.

Lisker, L. (1984). On Reconciling monophthongal vowel percepts and continuously varying F patterns. Haskins Laboratories: *Status Report on Speech research* SR-79/80.

Maeda, S. and Honda, K. (1994). From EMG to formant patterns of vowels: the implication of vowel spaces. *Phonetica*, 51, 17-29.

Mack, M. and Blumstein, S. (1983). Further evidence of acoustic invariance in speech production: the stop-glide contrast. *Journal of the Acoustical Society of America*, 73, 1739-1750.

Macneilage, P. F. (1970). Motor control of serial ordering of speech. *Psychological Review*, 182-196.

MacNeilage, P. and DeClerk, J. (1969). On the Motor Control of Coarticulation in CVC Monosyllables. *Journal of the Acoustical Society of America*, 45 (5), 1217-1233.

Magen, H. (1984). Vowel-to-vowel coarticulation in English and Japanese. *Journal of the Acoustical Society of America*, 75, Suppl. 1: S11.

Magen, H. (1989). An acoustic study of vowel-to-vowel coarticulation in English. *PhD thesis*, Yale.

Manuel, S. Y. and Krakow, R. A. (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. *Status Report on Speech Research*, Sr-77/78, Haskins Laboratories.

Manuel, S. Y. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different langauages. *Journal of the Acoustical Society of America*, 88 (3), 1286-1298.

Mermelstein, P. (1978). Difference limens for formant frequencies of steady-state and consonant-bound vowels. *Journal of the Acoustical Society of America*, 63 (2), 572-580.

Miller, J. E. and Fujimura, O. (1982). Graphic displays of combined presentations of acoustic and articulatory information. *The Bell System Technical Journal*, 61, 799-810.

Miller, J. D. (1989). Auditory-perceptual interpretation of the vowel. *Journal of the Acoustical Society of America*, 85 (5), 2114-2134.

Moll, K. L. and Daniloff, R. D. (1971). Investigation of the timing of velar movements during speech. *Journal of the Acoustical Society of America,* 50, 678-685.

Nearey, T. M. (1978). *Phonetic Feature Systems for Vowels.* Indiana University Linguistics Club, Bloomington, IN.

Nearey, T. M. (1980). On the physical interpretation of vowel quality: cinefluorographic and acoustic evidence. *Journal of Phonetics*, 8, 213-241.

Nearey, T. M. (1988). Applications of generalised linear modelling to vowel data. *Proceedings of the 1992 International Conference on Spoken Language Processing,* 583-586. Aix-en-Provence.

Nearey, T. M. (1989). Static, dynamic and relational properties in vowel perception. *Journal of the Acoustical Society of America*, 85 (5) 2088-2113.

Nolan, F. J. (1985). Idiosyncrasy in coarticulatory strategies. *Cambridge Papers in Phonetics and Experimental Linguistics* 4.

Nord, L. (1974). Vowel reduction-centralisation or contextual assimilation? *Speech Communication Seminar*, 149-154.

Nord, L. (1986). Acoustic studies of vowel reduction in Swedish. *KTH-QPSR,* Stockholm 4, 19-36.

Ohala, J. J. (1989). Sound change is drawn from a pool of synchronic variation. In: Breivik, L. E. and Jahr, E. H. (Eds.), Language Change: Contributions to the study of its causes, (Series: Trends in Linguistics, Studies and Monographs 43), Mouton de Gruyter, Berlin, 173-198.

Ohala, J. J. (1993). Coarticulation and Phonology. *Language and Speech*, 36, (2,3), 155-170.

Ohde, R. N. and Sharf, D. J. (1975). Coarticulatory effects of voiced stops on the reduction of acoustic vowel targets. *Journal of the Acoustical Society of America*, 58(4), 923-927.

Ohman, S. E. G. (1966). Coariculation in VCV utterances: spectrographic measurements. *Journal of the Acoustical Society of America*, 39, 151-168.

Ohman, S. E. G. (1967). Numerical model of coarticulation. *Journal of the Acoustical Society of America*, 4, 310-320.

Perkell, J. S. and Nelson, W. L. (1985). Variability in the production of the vowels /i/ and /a/. *Journal of the Acoustical Society of America*, 77 (5), 1889-1895.

Peterson, G. E. and Barney, H. L. (1952). Control methods used in the study of the vowels. *Journal of the Acoustical Society of America*, 24 (2), 175-184.

Peterson, G. E. and Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America,* 32(6), 693-703.

Pierrehumbert, J. (1980). The phonology and phonetics of English intonation. *Ph.D. thesis*, MIT.

Pierrehumbert, J. and Beckman, M. E. (1988). Japanese Tone Structure. *The MIT Press*, Cambridge, Massachusetts.

Pisoni, D. (1980). Variability of vowel formant frequencies and the quantal theory of speech. *Phonetica*, 37, 285-305.

Port, R. (1981). Linguistic timing factors in combination. *Journal of the Acoustical Society of America*, 69(1), 262-274.

Potter, R. K. and Steinberg, J. C. (1950). Toward the specification of speech. *Journal of the Acoustical Society of America*, 22(6), 807-820.

Recasens, D. (1984). V-to-C coarticulation in Catalan VCV sequences: an articulatory and acoustical study. *Journal of Phonetics*, 12, 61-73.

Recasens, D. (1986). An acoustic analysis of V-to-C and V-to-V coarticulatory effects on Catalan and Spanish VCV sequences. *Status Report on Speech Research* Sr-86/87. Haskins Laboratories.

Recasens, D. (1991). An electropalatographic and acoustic study of consonant-to-vowel coarticulation. *Journal of Phonetics*, 19, 177-192.

Rischel, J. (1990). Formal Linguistics and Real Speech. *Proceedings of the ETRW 'Phonetics and Phonology of Speaking Styles'*, 6, 1-9.

Rubin, P. Baer, T. and Mermelstein, P. (1981). An articulatory synthesizer for perceptual research, *Journal of the Acoustical Society of America*, 70, 321-328.

Saltzman, E. C. (1986) Task dynamic coordination of speech articulators: A preliminary model. In: Heuer, H. and Fromm, C. (Eds.), *Generation and Modulation* *of Action Patterns (Experimental Brain research series* 15). Springer-Verlag: New York.

Saltzman E. C. and Kelso, J. A. S. (1987). Skilled actions: A task dynamic approach. *Psychological Review*, 94, 84-106.

Saltzman, E. C., Goldstein, L., Browman, C. and Rubin, P. (1988). Modeling speech production using dynamic gestural structures. *Journal of the Acoustical Society of America*, 84, S146 (Abstract).

Schouten, M. E. and Pols, L. C. W. (1979). Vowel segments in consonantal contexts: a spectral study of coarticulation-part 1. *Journal of Phonetics*, 7, 1-23.

Sharf, D. J. and Ohde, R. N. (1975). Coarticulatory effects of voiced stops on the reduction of acoustic vowel targets. *Journal of the Acoustical Society of America*, 58(4), 923-927.

Sharf, D. J. and Ohde, R. N. (1981). Physiological, acoustic and perceptual aspects of coarticulation: Implications for the remediation of articulatory disorders. *Speech and Language: Advances in Basic Research and Practice*, 5, 153-247.

Shearme, J. N. and Holmes, J. N. (1962). An experimental study of the classification of sounds in continuous speech according to their distribution in the formant-1, formant-2 plane. *Proceedings of the 4th International Congress of Phonetic Sciences*, Mouton, The Hague, 233-240.

Shriberg, L. and Kent, R. (1982). *Clinical Phonetics*. MacMillan, New York.

Silverman, K. E. A. (1987). Structure and processing of the fundamental frequency contour. *Phd thesis*, University of Cambridge.

Sole, M. J. and Ohala, J. J. (1991). The phonological representation of reduced forms. Proceedings of the ESCA Workshop '*Phonetics and phonology of speaking styles: Reduction and elaboration in speech communication*', Barcelona, 49:1-49:5.

Stalhammar, U., Karlsson, I. and Fant, G. (1973). 'Contextual effects on vowel nuclei', *KTH Speech Transmission Laboratory Quarterly Progress and Status Report*, 4, Stockholm, 4, 1-18.

Steriade, D. (1987). Redundant values. *Chicago Linguistic Society*, 23, 339-362.

Stevens, K. N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. *Human Communication: A Unified View*. In: Denes, P. B. and David, J. R. E. E. (Eds.), 51-66. McGraw-Hill, New York.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3-45.

Stevens, K. N. and Blumstein, S. E. (1979). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 64, 1358-68.

Stevens, K. N. and Blumstein, S. E. (1981). The search for invariant acoustic correlates of phonetic features. In: Eimas, P. D. and Miller, J. L. (Eds.). Perspectives on the study of speech. New Jersey: Erlbaum.

Stevens, K. N. and House, A. S. (1955). Development of a quantitative description of vowel articulation. *Journal of the Acoustical Society of America*, 27(3), 111-128.

Stevens, K. N. and House, A. S. (1956). Studies of formant transitions using a vocal tract analog. *Journal of the Acoustical Society of America*, 28(4), 578-585.

Stevens, K. N. and House, A. S. (1963). Perturbation of vowel articulations by consonantal context: an acoustical study. *Journal of Speech and Hearing Research*, 6, 111-128.

Stevens, K. N., House, A. S. and Paul, A. P. (1966). Acoustical description of Syllabic nuclei: an interpretation in terms of a dynamic model of articulation. *Journal of the Acoustical Society of America*, 40, 123-132.

Strange, W. (1989a). Dynamic specification of coarticulated vowels spoken in a sentence context. *Journal of the Acoustical Society of America*, 85(5), 2135-2153.

Strange, W. (1989b). Evolving theories of vowel perception. *Journal of the Acoustical Society of America*, 85(5), 2081-2087.

Strange, W., Jenkins, J. J. and Johnson, T. L. (1983). Dynamic specification of coarticulated vowels. *Journal of the Acoustical Society of America*, 74(3), 695-705.

Strange, W., Verbrugge, R. R., Shankweiler, D. P. and Edman, T. R. (1976). Consonant environment specifies vowel identity. *Journal of the Acoustical Society of America*, 60, 213-224.

Summers, W. V. (1987). Effects of stress and final-consonant voicing on vowel production: articulatory and acoustic analyses. *Journal of the Acoustical Society of America*, 82(3), 847-863.

Sussman, H. M. (1990a). Acoustic correlates of the front/back vowel distinction: a comparison of transition onset vs steady state. *Journal of the Acoustical Society of America*, 88(1), 87-96.

Sussman, H. M. (1990b). An investigation of locus equations as a Source of Relational Invariance for Stop Place Categorisation. *Proceedings of the ETRW ' Phonetics and Phonology of Speaking Styles'*, 50, 1-12.

Sussman, H. M. (1994). The Phonological reality of locus equations across manner class distinctions: Preliminary Observations. *Phonetica*, 51, 119-131.

Sussman, H. M., McCaffrey, H. A. and Matthews, S. A. (1991). An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of the Acoustical Society of America*, 90(3), 1309-1325.

Syrdal, A. K. (1985). Aspects of a model of the auditory peripheral representation of English vowels. *Speech Communication*, 4, 121-135.

Syrdal, A. K. and Gopal, H. A. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America*, 79, 1086-1100.

Syrdal, A. K. and Steele, S. A. (1985). Vowel F1 as a function of speaker fundamental frequency. *Journal of the Acoustical Society of America*, Suppl. 1 78 (S56), 121-135.

Thompson, H. S. and Laver, J. D. (1987). The Alvey Speech Demonstrator-Architecture, Methodology and progress to Date, Speech Tec. 87'. 15-18.

Tiffany, W. R. (1959). Nonrandom sources of variation in vowel quality. *Journal of Speech and Hearing Research*, 2, 305-317.

Tokoma, S. (1993). Some arguments on vowel formant shift. UCL Working Papers, 235-253.

Traunmuller, H. (1981). Perceptual dimension of openess in vowels. *Journal of the Acoustical Society of America*, 69, 1465-1475.

Tuller, B., Harris, K. S. and Scott-Kelso, J. A. (1982). Stress and rate: differential transformations of articulation. *Journal of the Acoustical Society of America*, 71(6), 1534-1543.

Umeda, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America*, 58, 434-445.

Van Bergem, D. R. (1988). The first step to a better understanding of vowel reduction. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*, 12, 61-75.

Van Bergem, D. R. (1990). Acoustic and lexical vowel reduction. In *Proceedings of the ETRW Phonetics and Phonology of Speaking Styles*. Institute of Phonetic Sciences, University of Amsterdam.

Van Bergem, D. R. (1993). Acoustic vowel reduction as a function of sentence accent, word stress and word class. *Speech Communication*, 12, 1-23.

Van Bergem, D. R. (1994). A model of coarticulatory effects on the schwa. *Speech Communication* 14, 143-162.

Van Bergem, D. R. (1995). Acoustic and lexical vowel reduction. *PhD thesis*, University of Amsterdam.

Van Son, R. J. J. H. and Pols, L. C. W. (1990). Formant frequencies of Dutch vowels in a text read at normal and fast rate. *Journal of the Acoustical Society of America*, 88(4), 1683-1693.

Van Son, R. J. J. H. (1993). Spectro-temporal Features of Vowel Segments. *Ph.D thesis*, University of Amsterdam.

Veatch, T. C. (1991). English Vowels: their surface phonology and phonetic implementation in vernacular dialects. *PhD thesis*, University of Pennsylvania.

Verbrugge, R. R. and Shankweiler, D. P. (1977). Prosodic Information for vowel identity. *Journal of the Acoustical Society*, 61, Suppl. No. 1, S39.

Watson, G. S. (1988). APS User Guide, *CSTR Internal Document.*

Wells, J. C. (1962). A Study of formants of the pure vowels in British English. M.A. Thesis, University of London in Fry, 1979.

Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.